

SUMMARIES

OM protein - nucleic search, using frame_plus_p2n model
Run on: December 14, 2002, 18:11:26 ; Search time 139.234 Seconds
(without alignments)
3202.482 Million cell updates/sec

Title: US-09-782-953-3
Perfect score: 1048
Sequence: 1 MEEVDQLPNSATACHLDP.....RPKPKIOTRREPTPIHLS 198

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 112599159 residues
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DBV=xlh
-Q=/cgn2_1/USPTO.spool/US09782953/runat_11122002_114428_17330/app_query.fasta_1.1173
-DB=N Geneseq 101002 -QPMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFWT=ptn -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09782953 -CGN1_1_220 @runat_11122002_114428_17330 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -JONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 101002.*
1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1048	100.0	597	24	Human MCIP associa
2	1002.5	95.7	2212	24	Human MCIP associa
3	957.5	91.4	2358	21	Lung cancer associ
4	920.5	87.8	597	24	Mouse MCIP associa
5	872.5	83.3	2331	24	Human MCIP associa
6	859.5	82.0	2348	24	Human DNA sequence
7	847.5	80.9	2173	24	Gene #3341 used to
8	844.5	80.6	2174	20	Human DSCR1 coding
9	749.5	71.5	599	24	Human MCIP associa
10	620.5	59.2	934	22	Nucleotide sequenc
11	620.5	59.2	3159	24	Human Down syndrom
12	613.5	58.5	584	24	Mouse MCIP associa
13	596.5	56.9	3184	24	Human MCIP associa
14	596.5	56.9	3184	24	Colon adenocarcino
15	596.5	56.9	3184	24	Lung cancer relate
16	596	56.9	828	24	ABL65187
17	595	56.8	720	24	Human MCIP associa
18	539	51.4	615	24	Human MCIP associa
19	509.5	48.6	1021	24	Rat Down syndrome
20	439	41.9	442	21	Human secreted pro
21	379.5	36.2	1820	23	Drosophila melanog
22	379.5	36.2	12550	23	Drosophila melanog
23	285.5	27.2	412	22	Human foetal liver
24	285.5	27.2	412	22	Human brain expres
25	285.5	27.2	412	22	Human brain expres
26	285.5	27.2	412	22	Probe #4382 for ge
27	285.5	27.2	412	22	Probe #4507 used t
28	285.5	27.2	412	22	Human genome-deriv
29	285.5	27.2	412	24	Human breast cell
30	285.5	27.2	446	22	Human foetal liver
31	285.5	27.2	446	22	Probe #2064 for ge
32	285.5	27.2	446	22	Human brain expres
33	285.5	27.2	446	22	Human bone marrow
34	285.5	27.2	446	22	Probe #2081 for ge
35	285.5	27.2	446	22	Probe #2179 used t
36	285.5	27.2	446	22	Probe #2052 used t
37	285.5	27.2	446	22	Human genome-deriv
38	266.5	25.4	486	22	Human breast cell
39	266.5	25.4	486	22	Human foetal liver
40	266.5	25.4	486	22	Probe #809 for gen
41	266.5	25.4	486	22	Human brain expres
42	266.5	25.4	486	22	Human bone marrow
43	266.5	25.4	486	22	Probe #836 for gen
44	266.5	25.4	486	22	Probe #849 used to
45	266.5	25.4	486	22	Probe #816 used to

ALIGNMENTS

RESULT 1
AAD30152
ID AAD30152 standard; DNA; 597 BP.

AC AAD30152;
XX
XX 17-MAY-2002 (first entry)
DT
DE Human MCIP associated DNA #2.
XX
XX Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;
KW heart failure; cardiomyopathy; heart disease; human; gene; ds.
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FT CDS 1..597
FT /tag= a

```

FT      /product= "Human MCIP associated protein #2"
XX      MO200204491-A2.
XX      17-JAN-2002.
XX      06-JUL-2001; 2001WO-US21662.
XX      07-JUL-2000; 2000US-216601P.
XX      13-FEB-2001; 2001US-0782953.
XX      (TEXA ) UNIV TEXAS SYSTEM.
XX      (WILL/) WILLIAMS S R.
XX      (ROTH/) ROTHERMEL B.
XX      Williams SR, Rothermel B;
XX      MPI; 2002-179698/23.
XX      P-PSDB; AAET18911.
XX      Screening for modulators of muscle calcineurin interacting protein
XX      (MCIP) binding, expression or phosphorylation, useful for treating
XX      cardiac hypertrophy or heart failure, comprises mixing MCIP,
XX      calcineurin and a test compound -
XX      Claim 95; Page 147-148; 174pp; English.
XX      The invention relates to muscle calcineurin interacting proteins (MCIPs)
XX      and nucleic acid molecules encoding such proteins. MCIPs form a physical
XX      complex with the catalytic subunit of calcineurin and increased levels
XX      of MCIPs correspond to a reduced ability of calcineurin to stimulate
XX      transcription of certain target genes. The invention also relates to
XX      methods for identifying modulators of MCIP binding, expression or
XX      phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin
XX      may be used for treating cardiac hypertrophy and heart failure.
XX      CC Antibodies to MCIP can be used in characterizing the MCIP content of
XX      healthy and diseased tissues and subsequently for determining the
XX      presence or absence of cardiomyopathy or as predictor of heart disease.
XX      The present sequence is human MCIP associated DNA.
XX      Sequence 597 BP; 160 A; 170 C; 151 G; 116 T; 0 other;
XX      Alignment Scores:
XX      Pred. No.: 1.57e-109 Length: 597
XX      Percent: 1048.00 Matches: 198
XX      Score: 100.00% Conservative: 0
XX      Best Local Similarity: 100.00% Mismatches: 0
XX      Query Match: 100.00% Indels: 0
XX      Gaps: 0
XX      US-09-782-953-3 (1-198) x AAD30152 (1-597)
OY      1 MetGUGUValAAspLeuGlnAspLeuProSerAlaThrIleAlaCySHsLeuAspPro 20
Db      1 ATGGAGGAGGAGTGAGATCTGCAGAGACCTGCCGAGCCCAATCGCTCCACCTGCAGACCCG 60
OY      21 ArgValAPheValAspGlyLeuCyArgAlaIAspGlnSerLeuPheArgThrTyrAsp 40
Db      61 CGCGGTTCGAGAGCGGCTGTGCGCGGCCCAATTATTAATCCCTCTTCAGAAACATATGAC 120
OY      41 LysAspThrThrPheGlnTyrPheIysSerPheIysArgValArgIleAsnPheSerAsn 60
Db      121 AAGGACACACACCTCCAGTATTTTAAGAGCTTCAAACGTGCGGATTAACCTTCAGAAC 180
OY      61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuIleIysThrGlnPheLeuGlyIlyGlu 80
Db      181 CCTTATTTGCAGCCGATGCGAGCGCTGGCTGCACAAAGCCAGATTCCTCGGGAAAGGAA 240
OY      81 MetIysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProProAn 100
Db      241 ATGAAGGTGTATTTTGTGCTCAGACTTTACACATGTGGAAGATTCAACCTGGTCCGCCCAAT 300
OY      101 ProAspIysGlnPheLeuIleSerProProAlaSerProProValGlyTyrIlyGlnVal 120

```

Dd		301	CCGCAACAAAGTTCCTATCTCCTCCCTCGGGCTCTCCTCCGGTGGAACAAGA	*
Oy		121	GUAspAlaThr-ProValIleasnTyraPdeLeuLeuTyrAlaIleserHyseuglyPro	140
Dd		361	GAGAGGCCACCCTCCCTCATTAATTACGATCTTTATATGATGCATCTCCAAGCTGGGGCCA	420
Oy		141	GLYGLUeTYrGLIudenhIsalAaLThraPProthrProSerValValHisVal	160
Dd		421	GGAGAGAAATATGAACTGCATGCACGACGACCCCACTCCAGTGTGTGTCCACGTG	480
Oy		161	CyGglUseASPdGlInuansgLuaglLuaglLuagImetGuArgMetHyasArgPro	180
Dd		481	TGTGAGAGTGCACCAAGACATAGAGAGGAAGAAGAGATGAGATGAAGAAGAGACC	540
Oy		181	LysProLyAllelEgLeThrArgArgProGluTyrThrProIleHisLeuSer	198
Dd		541	AAGCCCAAATATCATCACAGACGGAGCCGGAGTACACACGATCCACTTAGC	594
RESULT 2				
AAD30156				
ID		AAD30156	standard; DNA; 2212 BP.	
XX				
AC		AAD30156;		
XX				
DT		17-MAY-2002	(first entry)	
XX				
DE			Human MCIP associated DNA #2.	
XX				
KM			Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;	
KW			heart failure; cardiomyopathy; heart disease; human; gene; ds.	
XX				
OS			Homo sapiens.	
XX				
Key			Location/Qualifiers	
FH		25..618		
FT		/**tag= a		
CDS		/product= "Human MCIP associated protein #2"		
PN		WO200204491-A2.		
PD		17-JAN-2002.		
XX				
PF		06-JUL-2001; 2001WC-US21662.		
PR		07-JUL-2000; 2000US-216601P.		
ER		13-FEB-2001; 2001US-0782953.		
XX				
PA		(TEXA) UNIV TEXAS SYSTEM.		
PA		(WILL/) WILLIAMS S R.		
PA		(ROTH/) ROTHERMEL B.		
PI		Williams SR, Rothermel B;		
DR		WPI; 2002-179698/23.		
P-PSDB;		AAE18915.		
PT			Screening for modulators of muscle calcineurin interacting protein	
PT			(MCIP) binding, expression or phosphorylation, useful for treating	
PT			cardiac hypertrophy or heart failure, comprises mixing MCIP,	
PT			calcineurin and a test compound -	
XX				
PS			Example 1; Page 159-161; 174pp; English.	
XX				
CC			The invention relates to muscle calcineurin interacting proteins (MCIPs)	
CC			and nucleic acid molecules encoding such proteins. MCIPs form a physical	
CC			complex with the catalytic subunit of calcineurin and increased levels	
CC			of MCIPs correspond to a reduced ability of calcineurin to stimulate	
CC			transcription of certain target genes. The invention also relates to	
CC			methods for identifying modulators of MCIP binding, expression or	
CC			phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin	
CC			may be used for treating cardiac hypertrophy and heart failure.	
CC			Antibodies to MCIP can be used in characterizing the MCIP content of	

CC healthy and diseased tissues and subsequently for determining the
 CC presence or absence of cardiomyopathy or as predictor of heart disease.
 CC The present sequence is human MCIP associated DNA.
 CC Note: this sequence has been described as murine MCIP splice variant in
 CC the specification, however the sequence seems to be a polynucleotide
 CC encoding a MCIP associated protein.

XX Sequence 2212 BP; 588 A; 455 C; 529 G; 640 T; 0 other;

Alignment Scores:
 Pred. No.: 1.38e-103 Length: 2212
 Score: 1002.50 Matches: 191
 Percent Similarity: 97.9% Conservative: 3
 Best Local Similarity: 96.46% Mismatches: 3
 Query Match: 95.66% Indels: 1
 DB: 24 Gaps: 1

US-09-782-953-3 (1-198) x AAB30156 (1-2212)

QY 1 MetGluGluValAspLeuGlnAspLeuProSerAlaThrIleAlaCysHisLeuAspPro 20
 DB 25 ATGAGGAGGTGGACCTGCGAGGACCTGCCAGCGCCACCATCGCTGTACCTGGACCG 84
 QY 21 ArgValPheValAspGlyLeuCysArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40
 DB 85 CGCGTGTTCGTGGACGGCTGTGCCGGCCAAATTTGAGTCCCTCTTTAGGACGTATGAC 144
 QY 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60
 DB 145 AAGGACATCACCTTTAGTATTTTAAAGAGCTTCAACGAGTCAAGATAAATTCACGCAAC 204
 QY 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
 DB 205 CCCTTCTCCGACGAGATGCCAGGCTCCAGTGCATGAAGACTGAGTTTCTGGGAAAGGAA 264
 QY 81 MetLysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProProAsn 100
 DB 265 ATGAAGTATATTTTGTCTGACACCTTACACATAGGAAGCTCACACCTGGCTCCGCCAAAT 324
 QY 101 ProAspLysGlnPheLeuIleSerProProAlaSerProValGlyTyrLysGlnVal 120
 DB 325 CCAGCAAGAGATTTCTGATCTCCCTCCCGCTCTCCGCGAGTGGATGGAACAAGTG 384
 QY 121 GluAspAlaThrProValIleAsnTyrAspLeuTyrAlaIleSerLysLeuGlyPro 140
 DB 385 GAAGATCGACCCAGTCATAAATATGATCTTATATGCTATCTCAAGCTGGGCGCA 444
 QY 141 GlyLysTyrTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160
 DB 445 GGGGAAAAGTATGAATTCACGCGAGCTGACACCACTCCAGCGTGGTCCATGTA 504
 QY 161 CysGluSerAspGlnGluAsnGluGluGluGluGluMetGluArgMetLysArgPro 180
 DB 505 TGTGAGAGTATCAAGAG---AAGGAGAGAGAGAGAAATGGAAAGANTGAGAGACCT 561
 QY 181 LysProLysIleIleGlnThrArgArgProGluTyrThrProIleHisLeuSer 198
 DB 562 AAGCAAAAATATCCAGACCGAGGCGGAGTACAGCGGATCCACCTCAGC 615

RESULT 3

AAFI8328

ID AAFI8328 standard; DNA; 2358 BP.

XX AC AAFI8328;

XX 14-MAR-2001 (first entry)

DT Lung cancer associated polynucleotide sequence SEQ ID 347.

DE Human; lung cancer associated protein; neuroprotective; cytoskeletal;

KW cardioactive; immunomodulatory; muscular active; vulnary;

KW gastrointestinal; nephrotropic; antiinfective; gynecological;

KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;

KW proliferative disorder; wound healing; infectious disease; ds.

XX Homo sapiens.

XX W0200055180-A2.

XX 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US05918.

XX 12-MAR-1999; 99US-0124270.

XX (HUMA-) HUMAN GENOME SCI INC.

XX (ROSE/) ROSEN C A.

XX Ruben SM;

XX WPI; 2000-587514/55.

XX P-FSDB; AAB58452.

XX Lung cancer associated gene sequences, referred to as lung cancer
 PT antigens, useful for treatment, prevention, and diagnosis of disorders
 PT such as lung cancer -
 XX Claim 1; Page 805-806; 1425pp; English.

XX Polynucleotide sequences AAFI7982 - AAFI8424 encode human lung cancer
 CC associated proteins represented in AAB58106 - AAB58548. Lung cancer
 CC associated proteins and polynucleotide sequences, their agonists, and
 CC antagonists may have neuroprotective; cytoskeletal; cardioactive;
 CC immunomodulatory; muscular active general; vulnary; gastrointestinal
 CC general; nephrotropic; antiinfective; gynecological; or antibacterial
 CC activity. The invention also includes antibodies specific for the
 CC protein or polynucleotide sequences. The lung cancer associated
 CC polynucleotide sequences may be used for detection of lung cancer,
 CC chromosome identification, as chromosome markers, and for numerous other
 CC diagnostic or research purposes. The proteins may be used to treat
 CC disorders such as neural, immune, muscular, reproductive,
 CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
 CC disorders. The proteins may also be used in the treatment of wounds and
 CC infectious diseases. Polynucleotide sequences AAFI8425 - AAFI8433 and
 CC peptide AAB58549 are used in the course of the invention for the
 CC identification and characterisation of the polynucleotide and protein
 CC sequences.

XX Sequence 2358 BP; 641 A; 482 C; 584 G; 647 T; 4 other;

Alignment Scores:

Pred. No.: 1.93e-98 Length: 2358

Score: 957.50 Matches: 188

Percent Similarity: 96.46% Conservative: 3

Best Local Similarity: 94.95% Mismatches: 6

Query Match: 91.36% Indels: 2

DB: 21 Gaps: 1

US-09-782-953-3 (1-198) x AAFI8328 (1-2358)

QY 1 MetGluGluValAspLeuGlnAspLeuProSerAlaThrIleAlaCysHisLeuAspPro 20
 DB 114 ATGAGGAGGTGGACCTGCGAGGACCTGCCAGCGCCACCATCGCTGTACCTGGACCG 173
 QY 21 ArgValPheValAspGlyLeuCysArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40
 DB 174 CGCGTGTTCGTGGACGGCTGTGCCGGCCAAATTTGAGTCCCTCTTTAGGACGTATGAC 233
 QY 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60
 DB 234 AAGGACATCACCTTTAGTATTTTAAAGAGCTTCAACGAGTCAAGATAAATTCACGCAAC 293
 QY 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
 DB 294 CCCTTCTCCGACGAGATGCCAGGCTCCAGCTGCATAAGACTGAGTTTCTGGGAAAGGAA 353

```

OY 81 MetlyseuTYrPhealInThleuHialleglyseSerHisleuAlaProProAsn 100
DB 354 ATGAAGTTATATTTTGTCTGACACCTTACACATGAAAGCTCACCTGGCTCCGCA-AAT 412
OY 101 ProAspLysGlnPheleuHialleSerProAlaSerProProValGlyTyrPlyGlnVal 120
DB 413 CCGAGACAGAGAGTTTGTGATCTCCCTCCCGCTCCGSCAGTGGATGGAACAAGTG 472
OY 121 GluAspAlaThrProValIleleuTYrAspLeuLeuTYrAlaIleSerLysleuGlyPro 140
DB 473 GAAGATCCGACCCGACATTAACCTATGATCTTATATGATCCATCTCCAAAGCTGGGGCCA 532
OY 141 GlyGluLeuTYrGluLeuHialaIaIaThrAspProThrProSerValValIleHialVal 160
DB 533 GGGGAAAGATGATGATGACAGCAGCAGCTGACACCTCCGCGGTGGTCCATGTA 592
OY 161 CysGluSerAspGlnGluAsnGluGluGluGluGluGluMetGluArgMetLysArgPro 180
DB 593 TGTAGAGTGATCAAGG---AAGAGGAAAGAAAGAGAAATGAAAGATGAGAGACCT 649
OY 181 LysProLysIleIleleuInThraGargProGluTYrThrProIleHisLeuSer 198
DB 650 AAGCCAAAATTAATCCAGACGAGGCGGAGTACACGCCGATCCACCTCAGC 703

RESULT 4
AAD30153
ID AAD30153 standard; DNA: 597 BP.
XX
AC AAD30153;
XX
DT 17-MAY-2002 (first entry)
XX
DE Mouse MCIP associated DNA #3.
XX
KM Muscle calcineurin interacting protein; MCIP, cardiac hypertrophy;
XX heart failure; cardiomyopathy; heart disease; mouse; gene; db.
XX
OS Mus musculus.
XX
PH Key Location/Qualifiers
FT 1..597
FT CDS /tag= a
FT /product= "Mouse MCIP associated protein #3"
XX
PN MO200204491-A2.
XX
PD 17-JAN-2002.
XX
PF 06-JUL-2001; 2001WO-US21662.
XX
PR 07-JUL-2000; 2000US-216601P.
PR 13-FEB-2001; 2001US-0782953.
XX
PA (TEXA ) UNIV TEXAS SYSTEM.
PA (WILL/) WILLIAMS S R.
PA (ROTH/) ROTHERMEL B.
XX
PI Williams SR, Rothermel B;
XX
DR MPI, 2002-179698/23.
DR P-PSDB; AAE18912.
XX
PT Screening for modulators of muscle calcineurin interacting protein
PT (MCIP) binding, expression or phosphorylation, useful for treating
PT cardiac hypertrophy or heart failure, comprises mixing MCIP,
PT calcineurin and a test compound -
XX
PS Discloure; Page 150-151; 174pp; English.
XX
CC The invention relates to muscle calcineurin interacting proteins (MCIPs)
CC and nucleic acid molecules encoding such proteins. MCIPs form a physical
CC complex with the catalytic subunit of calcineurin and increased levels
CC of MCIPs correspond to a reduced ability of calcineurin to stimulate

```

```

CC transcripition of certain target genes. The invention also relates to
CC methods for identifying modulators of MCIP binding, expression or
CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin
CC may be used for treating cardiac hypertrophy and heart failure.
CC Antibodies to MCIP can be used in characterizing the MCIP content of
CC healthy and diseased tissues and subsequently for determining the
CC presence or absence of cardiomyopathy or as predictor of heart disease.
CC The present sequence is mouse MCIP associated DNA.
CC Note: This sequence has been described as human MCIP3 encoding DNA
CC in the specification, however the sequence seems to be a polynucleotide
CC encoding a MCIP associated protein.
XX
SQ Sequence 597 BP, 170 A; 156 C; 142 G; 129 T; 0 other;

Alignment Scores:
Pred. No.: 4,63e-95 Length: 597
Score: 920.50 Matches: 178
Percent Similarity: 92.42% Conservative: 5
Best Local Similarity: 89.90% Mismatches: 12
Query Match: 87.83% Indels: 3
DB: 24 Gaps: 1

US-09-782-953-3 (1-198) x AAD30153 (1-597)
OY 4 ValAspLeuGlnAspLeuPro-----SerAlaThrIleAlaCysHisLeuAspPro 20
DB 1 ATGAGTTTATAGGACCTTACGTAACATTTTAGCTCCCGATGCTGTGTGGCAACAGAT 60
OY 21 ArgValPheValAspGlyLeuCYaArgAlaLysPheGluSerLeuPheAlaGlyThrAsp 40
DB 61 GATGCTTCAGCGAAAGTAGAGACCAAGCGCCAAATTTAAATCCCTTCAGAACATATGAC 120
OY 41 LysAspThrThrPheGlnTYrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60
DB 121 AAGAGACACACCTTCCAGTATTTAAGAGCTTCAAAGCTGCCAGTAACCTTCAGAAC 180
OY 61 ProLeuSerAlaIaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
DB 181 CCTTATCTGCAGCGGATGCGCAGGCTCGGCTGCACAAGCCGATTCCTGGGGAAGAA 240
OY 81 MetLysLeuTYrPhealInThleuHialleglyseSerHisleuAlaProProAsn 100
DB 241 ATGAAGTTGATTTTCTCTCAAGCTTTTACACATGAAAGTTCAACCTGCTCCGCCAAT 300
OY 101 ProAspLysGlnPheleuHialleSerProAlaSerProProValGlyTyrPlyGlnVal 120
DB 301 CCGCAAAACAGTTCCTCATCTCCCTCCGCGCTCTCTCCCTGGCTGGAAACAAGTA 360
OY 121 GluAspAlaThrProValIleleuTYrAspLeuLeuTYrAlaIleSerLysleuGlyPro 140
DB 361 GAAGATGCCACCCCGTCATTAATTAAGATCTTTATATCCCATCTCCAAAGCTGGGGCCA 420
OY 141 GlyGluLeuTYrGluLeuHialaIaIaThrAspProThrProSerValValIleHialVal 160
DB 421 GGAAGAGAGTATGATGATGCTGACGACAGACCAACCTCCAGTGtGGTCCACAGTG 480
OY 161 CysGluSerAspGlnGluAsnGluGluGluGluGluGluMetGluArgMetLysArgPro 180
DB 481 TGTAGAGTGAACCAAGAGATGAGAGAGAAAGAGATGAGAGAAATGAGAGAGAGCC 540
OY 181 LysProLysIleIleleuInThraGargProGluTYrThrProIleHisLeuSer 198
DB 541 AAGCCAAAATTAATCCAGACGAGAGCGGAGTACACACCGATCCACCTTAC 594

RESULT 5
AAD30155
ID AAD30155 standard; DNA: 2331 BP.
XX
AC AAD30155;
XX
DT 17-MAY-2002 (first entry)
XX
DE Human MCIP associated DNA #1.

```

XX Musclé calcineurin interacting protein; MCIP; cardiac hypertrophy;
KW heart failure; cardiomyopathy; heart disease; human; gene; ds.
XX Homo sapiens.
XX Key Location/Qualifiers
FT 144..737
FT CDS /tag= a
FT /product= "Human MCIP associated protein #1"
XX WO200204491-A2.
XX 17-JAN-2002.
XX 06-JUL-2001; 2001WO-US21662.
XX 07-JUL-2000; 2000US-216601P.
XX 13-FEB-2001; 2001US-0782953.
XX (TEXA) UNIV TEXAS SYSTEM.
XX (WILL/) WILLIAMS S R.
XX (ROTH/) ROTHERMEL B.
XX Williams SR, Rothermel B;
XX WPI; 2002-179698/23.
XX P-PSDB; AAE18914.
XX Screening for modulators of muscle calcineurin interacting protein
XX (MCIP) binding, expression or phosphorylation, useful for treating
XX cardiac hypertrophy or heart failure, comprises mixing MCIP,
XX calcineurin and a test compound -
XX Claim 72; Page 155-157; 174pp; English.
XX The invention relates to muscle calcineurin interacting proteins (MCIPs)
XX and nucleic acid molecules encoding such proteins. MCIPs form a physical
XX complex with the catalytic subunit of calcineurin and increased levels
XX of MCIPs correspond to a reduced ability of calcineurin to stimulate
XX transcription of certain target genes. The invention also relates to
XX methods for identifying modulators of MCIP binding, expression or
XX phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin
XX may be used for treating cardiac hypertrophy and heart failure.
XX Antibodies to MCIP can be used in characterising the MCIP content of
XX healthy and diseased tissues and subsequently for determining the
XX presence or absence of cardiomyopathy or as predictor of heart disease.
XX The present sequence is human MCIP associated DNA.
XX Note: This sequence has been described as a promoter in claim 72 of
XX the specification, however the sequence seems to be a polynucleotide
XX encoding a MCIP associated protein.
XX SQ Sequence 2331 BP; 630 A; 470 C; 547 G; 684 T; 0 other;
Alignment Scores:
Pred. No.: 8.41e-89 Length: 2331
Score: 872.50 Matches: 168
Percent Similarity: 93.09% Conservative: 7
Best Local Similarity: 89.36% Mismatches: 12
Query Match: 83.25% Indels: 1
DB: 24 Gaps: 1
US-09-782-953-3 (1-198) x AAD30155 (1-2331)
Qy 11 SerAlaThrIleAlaCysHisLeuAspProArgValPheValAspGlyLeuCysArgAla 30
Db 174 AGCTCCCTGGATTGCTGTGCGCAACAGTAGTGATATCTTCAGCGAAAGTGAAACCGAGGCC 233
Qy 31 LysPheGluSerLeuPheArgThrTyrAspLysAspThrThrPheGlnTyrPheLysSer 50
Db 234 AATTTAGTCCCTCTTTAGGACGTATGACGACATCACCTTTCAGTATTTTAAAGGC 293
Qy 51 PheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAspAlaArgLeuArg 70

Db 294 TTCAAAACGAGTCAGATAAACTTTCAGCAACCCCTTCTCCGAGCAGATGCCAGGCTCCAG 353
Qy 71 LeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyrPheAlaGlnThrLeuHis 90
Db 354 CTGCATAAGACTGAGTTCTTCGGAAAGGAATGAAGTTATATTTTGTCTCAGACCTTACAC 413
Qy 91 IleGlySerSerHisLeuAlaProProAsnProAspLysGlnPheLeuLysSerProPro 110
Db 414 ATAGGAAGCTCACACCTGGCTCCGCCAAATCCAGACAAGCAGTTCTTGATCTCCCTCCC 473
Qy 111 AlaSerProProValGlyTyrLysGlnValGluAspAlaThrProValIleAsnTyrAsp 130
Db 474 GCCTCTCCGCCAGTGGGATGGAAACAGTGGAAGATGCGACCCAGTCATATAACTATGAT 533
Qy 131 LeuLeuTyrAlaIleSerLysLeuGlyProGlyGluLysTyrGluLeuHisAlaAlaThr 150
Db 534 CTCATTATATGCCATCTCCAAGCTGGGCGCAGGGGNAAGTATGAATTGCACGCGGACT 593
Qy 151 AspProThrProSerValValHisValCysGluSerAspGlnGluAsnGluGluGlu 170
Db 594 GACACCACTCCCGAGCTGGTGGTCCATGTATGTGAGAGTGATCAAGAG--AAGGAGGAA 650
Qy 171 GluGluGluMetGluArgMetLysArgProLysProLysIleIleGlnThrArgArgPro 190
Db 651 GAAAGAGGAATGGAAGAAATGAGGAGACCTAAGCCAAAATTTATCCAGACCAGGAGCGCG 710
Qy 191 GluTyrThrProIleHisLeuSer 198
Db 711 GAGTACACCGCATCCACCTCAGC 734
RESULT 6
AAS94805
ID AAS94805 standard; DNA; 2348 BP.
XX AC AAS94805;
XX 14-FEB-2002 (first entry)
DT Human DNA sequence #60 expressed during foam cell differentiation.
DE Human; foam cell differentiation; atherosclerosis; cerebral stroke;
KW cardiovascular disorder; coronary artery disease; gene therapy; ds.
XX Homo sapiens.
XX WO200177389-A2.
XX 18-OCT-2001.
XX 04-APR-2001; 2001WO-US111128.
XX 05-APR-2000; 2000US-195106P.
XX (INCY-) INCYTE GENOMICS INC.
XX Shiffman D, Somogyi R, Lawn R, Seilhamer JJ, Porter GJ, Mikita T;
PI Tai J;
XX WPI; 2002-010925/01.
XX Composition useful for diagnosis of conditions, disorders or diseases
XX associated with atherosclerosis, comprises several polynucleotides that
XX are differentially expressed in foam cell development -
XX Claim 1; Page 112-113; 315pp; English.
XX The present invention relates to the isolation of human polynucleotide
XX sequences that are differentially expressed during foam cell
XX differentiation. The polynucleotide sequences of the invention or a
XX composition comprising these polynucleotides are useful as a high
XX throughput method for detecting altered expression of one or more
XX polynucleotides in a sample. The polynucleotides can be used in the


```

Db 358 GATCTCTTATATGCCATCTCCAGCTGGGGCCAGGGGAAAGATGATGAATTGACCGCAGCG 417
Qy 150 ThrAspProThrProSerValValHisValCysGluSerAspGlnGluAsnGluGlu 169
Db 418 ACTGACACCACTCCAGCGTGGTGGTCCATGTATGTGAGAGTGATCAAGAG---AAGGAG 474
Qy 170 GluGluGluGluMetGluArgMetLysArgProLysProLysIleGlnThrArgArg 189
Db 475 GAGAAGAGGAATGGAAAGATGAGGAGACCTAAGCCAAAATTTATCCAGCAGGAGG 534
Qy 190 ProGluTyrThrProIleHisLeuSer 198
Db 535 CCGGAGTACACCGCGATCCACCTCAGC 561

RESULT 8
ID AAX01282
AA01282 standard; cDNA to mRNA; 2174 BP.
XX AAX01282;
XX AC AAX01282;
XX 09-APR-1999 (first entry)
XX Human DSCR1 coding sequence.
XX DSCR1; Down's Syndrome critical region 1; human; pathogenic abnormality;
XX Central Nervous System development; mental retardation; heart defect; ds.
XX Homo sapiens.
XX Key Location/Qualifiers
XX CDS 49..564
XX /*tag= a
XX polyA_signal 1541..1546
XX /*tag= b
XX polyA_signal 2132..2137
XX /*tag= c
XX US5869318-A.
XX 09-FEB-1999.
XX 07-JUN-1996; 96US-0665040.
XX 07-JUN-1995; 95ES-0001140.
XX (PALL/) PALLEJA X E.
XX Fuentes JJ, Palreja XE, Pritchard M;
XX WPI; 1999-152781/13.
XX P-FSDB; AAW73898.
XX DNA encoding foetal brain proteins - believed to be associated with
XX Down's syndrome
XX Claim 1; Column 15-18; 19pp; English.
XX This sequence is the Down's Syndrome critical region 1 (DSCR1) gene
XX of the invention. The DSCR1 gene was found to be located in the
XX q22.1-22.2 region of human chromosome 21. An increase in the transient
XX expression of DSCR1 mRNA in the brains of young rats, compared to
XX expression levels in the brains of adult rats, suggests an important role
XX for DSCR1 during the development of the Central Nervous System (CNS), and
XX that over expression of DSCR1 may be involved in pathogenic abnormalities
XX of mental retardation and/or heart defects as found in Down's syndrome
XX patients.
XX SQ Sequence 2174 BP; 605 A; 436 C; 502 G; 628 T; 3 other;
XX Alignment Scores:
XX Pred. No.: 1.15e-85 Length: 2174
XX Score: 844.50 Matches: 161

```

```

Percent Similarity: 97.63% Conservative: 4
Best Local Similarity: 95.27% Mismatches: 3
Query Match: 80.58% Indels: 1
DB: 20 Gaps: 1
US-09-782-953-3 (1-198) x AAX01282 (1-2174)
Qy 30 AlaLysPheGluSerLeuPheArgThrTyrAspLysAspThrThrPheGlnTyrPheLys 49
Db 58 GCCAAATTTGAGTCCCTCTTTAGGAGCGTATGACAAGGACATCACCTTTTCAGTATTTTAA 117
Qy 50 SerPheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAspAlaArgLeu 69
Db 118 AGCTTTCAAACGAGTCAGAAATAAAGCTTCAGCAACCCCTTCTCCGAGCAGATGCCAGGCTC 177
Qy 70 ArgLeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyrPheAlaGlnThrLeu 89
Db 178 CAGCTGCATAGACTGAGTTCTGGGAAAGGAATGAAGTTATATTTTGTCTCAGACCTTA 237
Qy 90 HisIleGlySerSerHisLeuAlaProAsnProAspLysGlnPheLeuLeuSerPro 109
Db 238 CACATAGGAAGCTCACACTGGCTCCGCCAAATCCAGACAAGCAGTTTCTGATCTCCCT 297
Qy 110 ProAlaSerProProValGlyTyrPlyGlnValGluAspAlaThrProValIleAsnTyr 129
Db 298 CCCGCTCTCTCCGCCAGTGGGATGGAACAACACTGGGAAGATGGACCCAGTCATAAATAT 357
Qy 130 AspLeuLeuTyrAlaIleSerLysLeuGlyProGlyGluLysTyrGluLeuHisAlaAla 149
Db 358 GATCTCTTATATGCCATCTCCAGCTGGGGCCAGGGGAAAGATGATGAATTCACCGCAGCG 417
Qy 150 ThrAspProThrProSerValValHisValCysGluSerAspGlnGluAsnGluGlu 169
Db 418 ACTGACACCACTCCAGCGTGGTGGTCCATGTATGTGAGAGTGATCAAGAG---AAGGAG 474
Qy 170 GluGluGluGluMetGluArgMetLysArgProLysProLysIleGlnThrArgArg 189
Db 475 GAAGAAGAGGAATGGAAAGATGAGGAGACCTAAGCCAAAATTTATCCAGCAGGAGG 534
Qy 190 ProGluTyrThrProIleHisLeuSer 198
Db 535 CCGGAGTACACCGCGATCCACCTCAGC 561

RESULT 9
AAX01282
ID AAD30151 standard; DNA; 599 BP.
XX AAD30151;
XX AC AAD30151;
XX DT 17-MAY-2002 (first entry)
XX DE Human MCIP associated DNA #1.
XX KW Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;
XX heart failure; cardiomyopathy; heart disease; human; ds.
XX OS Homo sapiens.
XX PN WO200204491-A2.
XX PD 17-JAN-2002.
XX PP 06-JUL-2001; 2001WO-US21662.
XX PR 07-JUL-2000; 2000US-216601P.
XX PR 13-FEB-2001; 2001US-0782953.
XX (TEXA ) UNIV TEXAS SYSTEM.
XX PA (WILL/) WILLIAMS S R.
XX PA (ROTH/) ROTHERMEL B.
XX PI Williams SR, Rothermel B;
XX

```

DR WPI; 2002-179698/23.
 XX Screening for modulators of muscle calcineurin interacting protein
 PT (MCIP) binding, expression or phosphorylation, useful for treating
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,
 PT calcineurin and a test compound -
 XX
 PS Disclosure; Page 147; 174pp; English.
 XX
 CC The invention relates to muscle calcineurin interacting proteins (MCIPs)
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical
 CC complex with the catalytic subunit of calcineurin and increased levels
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate
 CC transcription of certain target genes. The invention also relates to
 CC methods for identifying modulators of MCIP binding, expression or
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin
 CC may be used for treating cardiac hypertrophy and heart failure.
 CC Antibodies to MCIP can be used in characterizing the MCIP content of
 CC healthy and diseased tissues and subsequently for determining the
 CC presence or absence of cardiomyopathy or as predictor of heart disease.
 CC The present sequence is human MCIP associated DNA.
 XX
 SQ Sequence 599 BP; 164 A; 150 C; 149 G; 136 T; 0 other;
 XX
 Alignment Scores:
 Pred. No.: 1,18e-75 Length: 599
 Score: 749.50 Matches: 151
 Percent Similarity: 90.12% Conservative: 4
 Best Local Similarity: 87.79% Mismatches: 14
 Query Match: 71.52% Indels: 4
 DB: 24 Gaps: 1
 US-09-782-953-3 (1-198) x AAD30151 (1-599)
 QY 5 AspleuGlnAspleuPro-----SerAlaThrIleAlaCyHisLeuAspProArg 21
 DB 83 GATTATTAAGGACCTTATGATCAATTTTACCTCCGATTCCTGCTGGCAACACATGAT 142
 QY 22 ValPheValAspGlyLeuCybAArgAlaValPheGluSerLeuPheArgThrTyArgLys 41
 DB 143 GTCCTCAGCGAAGTGAAGACCGGGCCAAATTTGAATCCCTTCAGAAACATATGACAAG 202
 QY 42 AspThrThrPheGlnTyPheLysSerPheLysArgValArgIleAspPheSerAspPro 61
 DB 203 GACACACACCTTCCAGTATTTAGAGCTTCAACCGTCCGATTAACCTTCAGCAACCC 262
 QY 62 LeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheGluGlyLysGluMet 81
 DB 263 TTATCTGACGCGGATGCCAGCTGGCGGCGCAACAGCCGACTTCTGGGAGAGAAATG 322
 QY 82 LysLeuTyrrPheAlaGlnThrLeuHisLysGlySerSerHisLeuAlaProProAspPro 101
 DB 333 AAGTGTATTTTCTCAGACTTACACATAGAGAGTTCAACCTGCTCCGCCAAT-CCC 381
 QY 102 AspLysGlnPheLeuIleSerProPheAlaSerProProValGlyTyrLysGlnValGlu 121
 DB 382 GACAAACAGCTTCTCATCTCCCTCCGCCCTCTCTCCCGTTGGCTGGAAACAGTAGAA 441
 QY 122 AspAlaThrProValIleLeuAsnTyrrAspleuLeuTyrrAlaIleSerLysLeuGlyProGly 141
 DB 442 GATGCCACCCCGCTCATTAATTAACATCTTTATATGTCATCTCAACCTGGGGCAGAGA 501
 QY 142 GtLysTyrrGtLysHisAlaAlaThrAspProThrProSerValValHisValCys 161
 DB 502 GAGAAAGTATGAATGTCATGACGACGACGACAGACCACTCCACAGTGTGTGTCCACGTGTGT 561
 QY 162 GluSerAspGlnGluAsnGluGluGluGluGlu 173
 DB 562 GAGAGTCACCAAGAGATGAGAGAGAGAGAGAG 597
 RESULT 10
 AAF25338
 ID AAF25338 standard; cDNA; 934 BP.

XX
 AC AAF25338;
 XX
 XX 30-APR-2001 (first entry)
 DT
 XX
 DE Nucleotide sequence of a human detoxification protein.
 XX
 KW Human; detoxification protein; DEXT; cancer; leukemia; melanoma;
 KW adenocarcinoma; autoimmune disorder; inflammatory disorder;
 KW rheumatoid arthritis; asthma; atherosclerosis; autoimmune thyroiditis;
 KW psoriasis; ulcerative colitis; infection; cell proliferative disorder;
 KW actinic keratosis; arteriosclerosis; cirrhosis; hepatitis; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 20..724
 FT /*tag= a
 FT /*product= "detoxification protein"
 FT sig_peptide 20..100
 FT /*tag= b
 XX
 PN MO200104305-A2.
 XX
 PD 18-JAN-2001.
 XX
 PF 06-JUL-2000; 2000MC-US18509.
 XX
 PR 07-JUL-1999; 99US-0142678.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Tang YT, Yue H;
 XX
 DR WPI; 2001-147193/15.
 XX
 DR P-PSDB; AAB31788.
 XX
 PT New human detoxification protein and polynucleotide, useful for
 PT diagnosis, prevention and treatment of autoimmune/inflammatory
 PT disorders and cell proliferative disorders including cancer -
 XX
 PS Claim 5; Page 79; 79pp; English.
 XX
 CC The present sequence encodes a human detoxification polypeptide (DEXT).
 CC DEXT and its (ant)agonists are useful for preventing or treating
 CC disorders associated with decreased or increased expression or activity
 CC of DEXT. DEXT polypeptides are useful for screening compounds that
 CC specifically bind to DEXT and for identifying (ant)agonists.
 CC Diseases prevented, treated and diagnosed include cancers (e.g.
 CC leukemia, melanoma, adenocarcinoma, cancers of the brain, bone marrow,
 CC breast, kidney, liver, pancreas, prostate and uterus),
 CC autoimmune/inflammatory disorders (e.g. rheumatoid arthritis, asthma,
 CC atherosclerosis, autoimmune thyroiditis, psoriasis and ulcerative
 CC colitis), bacterial, fungal, parasitic infections and cell
 CC proliferative disorders (e.g. actinic keratosis, arteriosclerosis,
 CC cirrhosis and hepatitis). Anti-DEXT antibodies may be used as
 CC antagonists, as a targeting or delivery mechanism for bringing
 CC pharmaceutical agents into contact with cells or tissues expressing
 CC DEXT and for diagnosis of DEXT-related disorders.
 XX
 SQ Sequence 934 BP; 247 A; 244 C; 206 G; 237 T; 0 other;
 XX
 Alignment Scores:
 Pred. No.: 9.44e-61 Length: 934
 Score: 620.50 Matches: 121
 Percent Similarity: 73.71% Conservative: 22
 Best Local Similarity: 62.37% Mismatches: 42
 Query Match: 59.21% Indels: 9
 DB: 22 Gaps: 2
 US-09-782-953-3 (1-198) x AAF25338 (1-934)
 QY 5 AspleuGlnAspleuProSerAlaThrIleAlaCyHisLeuAspProArgValPheVal 24

```

Db 143 GACTTCAATGACCTCCCAACTCGTTGTTGGTGAATGTTCCACGAGTCAGTGTGTGAA 202
Qy 25 AspGlyLeuCyAspAlaLysPheGluSerLeuPheArgThrTyrAspLysAspThrThr 44
Db 203 GGAGAAGAGAGCAAGGAAATTTGAGGAGCTGTTTCGGACTTATGATGACTGTGTGACG 262
Qy 45 PheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsnProLeuSerAla 64
Db 263 TTCAGCTATTAAAGATTTTCAGACGTGTCGTATATAACTTCAGCAATCTAAATCTGCA 322
Qy 65 AlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyr 84
Db 323 GCCCGAGCTAGGATAGAGCTTCATGAACCAATTCAGAGGGAATAATTAAGCTCTAC 382
Qy 85 PheAlaGlnThrLeuHisLysLeGlySerSer-----HisLeuAlaProProAsn 100
Db 383 TTTCACAGCTTCAGACTCCAGAGACAGATGGAGACAACTGGCTCCACCCAG 442
Qy 101 ProAspLysGlnPheLeuLysSerProProAlaSerProProValGlyTyrLysGlnVal 120
Db 443 CCGGCCAAACAGTTCTCATCTCGCCCTCTCTCCACCTCTGGCTGGCAGCCATC 502
Qy 121 GluAspAlaThrProValIleAsnTyrAspLeuTyrAlaIleSerLysLeuGlyPro 140
Db 503 AAGATGCCAGCGCAGTCTCAATGATGACTCTCTATGCTGTGGCCAACTAGGACCA 562
Qy 141 GlyGlyLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160
Db 563 GGAGAGAAGTATGAGTCCATGAGGAGCTGAGTCCACCCCAAGTGTGTGTGACGTG 622
Qy 161 CysGluSerAspGlnGluAsnGluGluGluGluMetGluArgMetLysArgPro 180
Db 623 TGGCAGAGTGACATAGAGGAAGAGAGGCCAAAG-----ACTTCCCA 667
Qy 181 LysProLysIleLeuGlnThrArgArgProGluTyrThrPro 194
Db 668 AAGCCAAATAATCATCCAACTCGCGCTGTGGCTGCCACCC 709

RESULT 11
ID ABA91457
AC ABA91457;
XX
XX
XX 18-APR-2002 (first entry)
XX
XX Human Down syndrome critical region 1-like 1 protein cDNA.
XX
XX Down syndrome critical region 1-like 1; DSCR1L1 alpha; human;
XX Down syndrome; Alzheimer's disease; dementia; transgenic;
XX neuroprotective; nontropic; anticonvulsant; diagnosis;
XX gene therapy; gene; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 185..952
XX
XX /tag= a
XX /product= "DSCR1L1"
XX /transl_except= (pos:215..217, aa:Xaa)
XX /note= "Xaa = unknown"
XX
XX MO200204513-A2.
XX
XX
XX 17-JAN-2002.
XX
XX 11-JUL-2001; 2001WO-US21982.
XX
XX 11-JUL-2000; 2000US-0614474.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX

```

```

PI Loring JF, Tingley DW, Edwards CM, Streeter DG;
XX
XX WPI; 2002-164633/21.
XX P-PSDB; AAM50760.
XX
XX Novel Down syndrome critical region 1-like 1 protein and nucleic acid
XX encoding the protein useful for diagnosis and treatment of Alzheimer's
XX disease, Down syndrome and other forms of dementia -
XX
XX Claim 2; Page 46-47; 54pp; English.
XX
XX The present sequence is that of cDNA clone incyte ID No: 247500.5
XX encoding novel Down syndrome critical region 1-like 1 protein
XX (DSCR1L1 alpha, see AAM50760). Northern analysis indicated
XX expression of DSCR1L1 alpha in various libraries, with the highest
XX abundance in tissues from the nervous system, including tissues
XX associated with schizophrenia, Huntington's disease, epilepsy and
XX amyotrophic lateral sclerosis. An absence of DSCR1L1 expression
XX was observed in 7 of 8 libraries from subjects with Alzheimer's
XX disease. A claimed method for detecting differential expression of
XX a nucleic acid encoding DSCR1L1 alpha can be used to diagnose Down
XX syndrome, Alzheimer's disease and other forms of dementia. A probe
XX from such a nucleic acid is useful for identifying naturally
XX occurring molecules encoding DSCR1L1 alpha, allelic variants or
XX other molecules. The nucleic acids are useful for producing
XX transgenic cell lines or organisms which model human disorders.
XX They may also be used in gene therapy, and to screen for ligands
XX which specifically bind the nucleic acid molecule, such as
XX a DNA or RNA molecule, peptide nucleic acid, artificial chromosome,
XX peptide, transcription factor, repressor or regulatory molecule,
XX for use as a therapeutic.
XX
XX Sequence 3159 BP; 926 A; 657 C; 657 G; 915 T; 4 other;
XX
XX
XX Alignment Scores:
XX Pred. No.: 5,038-60 Length: 3159
XX Score: 620.50 Matches: 121
XX Percent Similarity: 73.71% Conservative: 22
XX Best Local Similarity: 62.37% Mismatches: 42
XX Query Match: 59.21% Indels: 9
XX DB: 24 Gaps: 2
XX
XX US-09-782-953-3 (1-198) x ABA91457 (1-3159)
XX
XX Qy 5 AspLeuGlnAspLeuProSerAlaThrIleAlaCysHisLeuAspProArgValPheVal 24
XX Db 371 GACTTCAATGACCTCCCAACTCGTTGTTGGTGAATGTTCCACGAGTGTGTGAA 430
XX
XX Qy 25 AspGlyLeuCyAspAlaLysPheGluSerLeuPheArgThrTyrAspLysAspThrThr 44
XX Db 431 GGAGAAGAGAGCAAGGAAATTTGAGGAGCTGTTTCGGACTTATGATGACTGTGTGACG 490
XX
XX Qy 45 PheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsnProLeuSerAla 64
XX Db 491 TTCAGCTATTAAAGATTTTCAGACGTGTCGTATATAACTTCAGCAATCTAAATCTGCA 550
XX
XX Qy 65 AlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyr 84
XX Db 551 GCCCGAGCTAGGATAGAGCTTCATGAACCAATTCAGAGGGAATAATTAAGCTCTAC 610
XX
XX Qy 85 PheAlaGlnThrLeuHisLysLeGlySerSer-----HisLeuAlaProProAsn 100
XX Db 611 TTTCACAGCTTCAGACTCCAGAGACAGATGGAGACAACTGGCTCCACCCAG 670
XX
XX Qy 101 ProAspLysGlnPheLeuLysSerProProAlaSerProProValGlyTyrLysGlnVal 120
XX Db 671 CCGGCCAAACAGTTCTCATCTCGCCCTCTCTCCACCTCTGGCTGGCAGCCATC 730
XX
XX Qy 121 GluAspAlaThrProValIleAsnTyrAspLeuTyrAlaIleSerLysLeuGlyPro 140
XX Db 731 AAGATGCCAGCGCAGTCTCAATGATGACTCTCTATGCTGTGGCCAACTAGGACCA 790
XX
XX Qy 141 GlyGlyLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160

```

D6			:		850
D6	791	GGAGGAAATGATGACCTCCATCGACGGACTAGTCACCCCAAGTGTGCACGTG			
OY	161	CyGgUSeRPePGLIGLuAbnGlUGluGlUGluGluMeGtGUARMeLyuArgPro			180
DB	851	TCCGACAGTGACATAGAGGAAGAAGAACGCCAAG-----ACTTCCCA			895
OY	181	LyBProLyvllleIeGlnThrArgArProGluTyThrPro			194
DB	896	AAGCCAAAATCATCTCAAACCTCGGCTCCTGCTGCCACC			937
RESULT 12					
ID	AAD30154	standard; DNA; 594 BP.			
XX	AAD30154				
XX	AAD30154;				
DT	17-MAY-2002	(first entry)			
XX					
DE	Mouse MCIP associated DNA #4.				
KW	Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;				
KX	heart failure; cardiomyopathy; heart disease; mouse; gene; da.				
OS	Mus musculus.				
XX					
FH	Location/Qualifiers				
FT	Key	1..594			
FT	CDS	/*arg= A			
FT		/product= "Mouse MCIP associated protein #4"			
PN	WO20020491-A2.				
PD	17-JAN-2002.				
PE	06-JUL-2001; 2001WO-US21662.				
XX					
PR	07-JUL-2000; 2000US-216601P.				
PR	13-FEB-2001; 2001US-0782953.				
PA	(TEXA) UNIV TEXAS SYSTEM.				
PA	(WILL/) WILLIAMS S R.				
XX	(ROTH/) ROTHERMEL B.				
P1	Williams SR, Rothermel B;				
DR	WPI: 2002-179698/23.				
DR	P-P5DB; AME18913.				
PT	Screening for modulators of muscle calcineurin interacting protein				
PT	(MCIP) binding, expression or phosphorylation, useful for treating				
PT	cardiac hypertrophy or heart failure, comprises mixing MCIP,				
PT	calcineurin and a test compound -				
PS	Disclosure; Page 153-154; 174pp; English.				
XX					
XX	The invention relates to muscle calcineurin interacting proteins (MCIPs)				
CC	and nucleic acid molecules encoding such proteins. MCIPs form a physical				
CC	complex with the catalytic subunit of calcineurin and increased levels				
CC	of MCIPs correspond to a reduced ability of calcineurin to stimulate				
CC	transcription of certain target genes. The invention also relates to				
CC	methods for identifying modulators of MCIP binding, expression or				
CC	phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin				
CC	may be used for treating cardiac hypertrophy and heart failure.				
CC	Antibodies to MCIP can be used in characterizing the MCIP content of				
CC	healthy and diseased tissues and subsequently for determining the				
CC	presence or absence of cardiomyopathy or as predictor of heart disease.				
CC	Note: This sequence has been described as mouse MCIP1 in the				
CC	specification, however the sequence seems to be a polynucleotide				
CC	encoding a MCIP associated protein.				
XX					
XX	Sequence 594 BP; 148 A; 165 G; 149 G; 132 T; 0 other;				

Alignment Scores:			
Pred. No.:	3,16e-60	Length:	594
Score:	613.50	Matches:	119
Percent Similarity:	75.00%	Conservative:	22
Best Local Similarity:	63.30%	Mismatches:	38
Query Match:	58.54%	Indels:	9
DB:	24	Gaps:	2

US-09-782-953-3 (1-198) x AAD30154 (1-594)

OY	11	SerAlaThrIleAlaCYenHLeuAspProArgValPheValAspGlyLeuCYArgAla	30
Db	31	TCCACTCGTCGCCTGTGTGGATGTGGAGTCTTTACCAATCGAGATTAAAGAA	90
OY	31	LysPheGluSerLeuPheArgThrTYrAspLysAphThrPheGlnTYrPheLysSer	50
Db	91	AAATTCCAGGAGCTGTTCCGAGCTTAATGATGATGATGAGCTTCCAGCTGTTTAAAGT	150
OY	51	PheLysArgValArgIleAsnPheserAsnProLeuSerAlaAlaAspAlaArgLeuArg	70
Db	151	TTCCACGCGGTTCGAATAATTTCAGCATCCTCCAATCTGCGAGCCGTCGGATGAG	210
OY	71	LeuHisLYSThrGluPheLeuGLYLeuGluMetCysLeuTYrPheAlaGlnThrLeuHis	90
Db	211	CTTCATGAGACTCAAGTTCAGAGGAGAAGACTAAACTTACTTGCCGCCAGGTCCAACC	270
OY	91	IleGlySerSer-----HisLeuAlaProFasnProAspLysGlnPheLeu	106
Db	271	CCAGAGACGATGAGACAACCTGCATTGGCACCTCCACAGCCTGCCAACAGTTCCTC	330
OY	107	IleSerProProAlaSerProProValGIYTrpLysGlnValGluAspAlaThrProVal	126
Db	331	AFTCAACCCCTCACTACCTCCCTCGTGGCGGAAGCTTACACGATGCCACACCACTC	390
OY	127	IleAsnTYrAspLeuLeuTYrAlaIleSerCysLeuGluYProGlyGluTYrTYrGluLeu	146
Db	391	CTCAACTATGACCTTCTTTATAGCTGTGTGGCCAACTTAGAGCAGAGAAATATAGACTG	456
OY	147	HisAlaAlaThrAspProThrProSerValValHisValCysGluSerAspGlnGlu	166
Db	451	CACGCTGGAACCTGAGTACACCCAGCGGTCTGTGATGTGTGACAGGACATGAG	516
OY	167	AsnGluGluGluGluGluGluMetCyluArgMetCyluArgProLysProLysIleIleGln	186
Db	511	GAGGAGGAGACCCAAAG-----ACTTCCCCCAAGCCAAAATACTATTACAG	555
OY	187	ThrArgArgProGluTYrThrPro	194
Db	556	ACCGGCGCTCCGGGCTTGCCACC	579

RESULT 13

AAD30157 standard; DNA; 3184 BP.

AAD30157;

17-MAY-2002 (first entry)

Human MCIP associated DNA #3.

Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;
heart failure; cardiomyopathy; heart disease; human; gene; ds.

Homo sapiens.

Key Location/Qualifiers

FT CDS 205..783

FT /tag= a

FT /product= "Human MCIP associated protein #3"

XX WO200204491-A2.

XX	(AVAL-) AVALON PHARM.
PA	
PI	Young PR, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
P1	Soppet DR, Weaver Z;
XX	
XX	WPI; 2002-188264/24.
DR	
PT	Screening for anti-neoplastic agent involves exposing cells to a
PT	chemical agent to be tested for anti-neoplastic activity, and
XX	determining a change in expression of a gene of a signature gene set -
PS	Claim 1; SEQ ID 105; 44pp; English.
XX	
CC	The present invention describes a method (M1) for screening for an
CC	anti-neoplastic agent. The method involves exposing cells to a chemical
CC	agent to be tested for anti-neoplastic activity, determining a change in
CC	expression of at least one gene (I) of a signature gene set, where (I)
CC	comprises a sequence (S) selected from 8447 sequences [given in ABU61664
CC	to ABU70110], or is at least 95% identical to (S), where a change in
CC	expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC	activity and can be used in gene therapy. M1 can be used for screening
CC	an anti-neoplastic agent, and can be used for producing a product which
CC	is the data collected with respect to the anti-neoplastic agent as a
CC	result of M1, and the data is sufficient to convey the chemical
CC	structure and/or properties of the agent. M1 can be used in the
CC	treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC	ovarian, kidney, prostate or pancreatic cancer.
CC	adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer
CC	infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC	carcinoma, papillary carcinoma and Wilms' tumor.
XX	
SO	Sequence 3184 BP; 921 A; 681 C; 657 G; 925 T; 0 other;
Alignment Scores:	
Pred. No.:	2,69e-57 Length: 3184
Score:	596.50 Matches: 117
Percent Similarity:	73.40% Conservative: 21
Best Local Similarity:	62.23% Mismatches: 41
Query Match:	56.92% Indels: 9
DB:	24 Gaps: 2
US-09-782-953-3 (1-198) x ABU61768 (1-3184)	
OY	11 SerAlatPrILleAlaCyAnHleuAnPRProAdgValAPheValAspGlyLeuCysArgAla 30
Dd	220 TCACCTGTGGTTCCTGTGTGGTGAGATGTGCAGAGCTTTTAACAATGAGAGATTAAAGA 279
OY	31 LysPhegiuSerLeuPheArgThrTyraRpyLysASPThrThPhesInTrPheIySer 50
Dd	280 AAATTTGGGGAGACTGTTTCGCACTTATGATGACTGTGACCGTCCAGCTATTTAAGAGT 339
OY	51 PheLyArGValArgIleAnPheSerAnPROLeuSerAlaAlaAspAlaArgLeuArg 70
Dd	340 TTCACACGGTCCGATAAATTCAGCAATTCCTAAATTCGACGCCGAGCTAGGAATGAG 399
OY	71 LeuHiVylrThGlupHeuGLyLysGluMelvLeuTyrrPheAlaglnThrLeuHis 90
Dd	400 CTTCATGAAAACCCAATTCAGAGGAAAAAATTAAACCTTACTTTCACACAGTTCAAGACT 459
OY	91 IleGIySerSer-----HisLeuAlaPRProAnPRroApnryLysGlnPheLeu 106
Dd	460 CCAGAGCAGATGAGACAAATCTGCACCTTGCGCTCCACCAGCTCCGAAACAGTTTCTC 519

Oy	107	IleSerProFoaIaSerPProFoValGlyTPbVgInValGIuSpAlaThrProVal	126
Db	520	ATCTCGCCCCCTTCTCCACCTGTAGTCGACGCCCATCAACAATGCCACGGCAGTC	579
Oy	127	IleAsnTYrAspLeuLeuTYrAlaIleSerLYsLeuGIYProGIYGIuTYrTYrGIuLeu	146
Db	580	CTCAACTATGAACCTCTCTATGCTGTGGCCAAACTAGAACGACGAGAGAAGATATGAGTC	639
Oy	147	HisaIalAthrAsPProThrProSeRvalValValHiseValCYsgIuSerAspInGIu	166
Db	640	CATGACGAGGACTGAGTCACACCACCAAGTGTCGTGCACGTGGCACAGTAGACATAGAG	699
Oy	167	AenGIuGIuGIuGIuGIuGIuMetGIuArgMeTYsArYProLYsProLYsIleIleGIu	186
Db	700	GAAAGAGAGGACCCCAAAG-----ACTTCCCANAAGCCMAATATCATCCA	744
Oy	187	ThirArgARProGIuTYrTYrPro	194
Db	745	ACTGGCGCTCCTGGCGCTGGCCACCC	768
RESULT 15			
ABL65187	ID	ABL65187 standard; DNA; 3184 BP.	
XX	AC	ABL65187;	
XX	DT	15-MAY-2002 (first entry)	
DE	XX	Lung cancer related gene sequence SEQ ID NO:3524.	
XX	XX	Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;	
KW	stomach; lung; prostate; pancreas; carcinoma; adenocarcinoma;		
KW	cystosctetic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;		
gene; ds.	XX		
OS	Homo sapiens.		
XX	PN	WO200194629-A2.	
XX	PD	13-DEC-2001.	
XX	PF	30-MAY-2001; 2001MO-US10838.	
XX	PR	05-JUN-2000; 2000US-209473P.	
PR	18-SEP-2000; 2000US-209531P.		
PR	18-SEP-2000; 2000US-233133P.		
PR	20-SEP-2000; 2000US-233617P.		
PR	20-SEP-2000; 2000US-234034P.		
PR	20-SEP-2000; 2000US-234052P.		
PR	22-SEP-2000; 2000US-234509P.		
PR	22-SEP-2000; 2000US-234567P.		
PR	25-SEP-2000; 2000US-234923P.		
PR	25-SEP-2000; 2000US-234924P.		
PR	25-SEP-2000; 2000US-235077P.		
PR	25-SEP-2000; 2000US-235082P.		
PR	25-SEP-2000; 2000US-235134P.		
PR	25-SEP-2000; 2000US-235280P.		
PR	26-SEP-2000; 2000US-235537P.		
PR	26-SEP-2000; 2000US-235538P.		
PR	27-SEP-2000; 2000US-235711P.		
PR	27-SEP-2000; 2000US-235720P.		
PR	27-SEP-2000; 2000US-235840P.		
PR	27-SEP-2000; 2000US-235863P.		
PR	28-SEP-2000; 2000US-236028P.		
PR	28-SEP-2000; 2000US-236032P.		
PR	28-SEP-2000; 2000US-236033P.		
PR	28-SEP-2000; 2000US-236034P.		
PR	28-SEP-2000; 2000US-236109P.		
PR	28-SEP-2000; 2000US-236111P.		
PR	29-SEP-2000; 2000US-236442P.		
PR	29-SEP-2000; 2000US-236891P.		

PR 02-OCT-2000; 2000US-237172P.
 PR 02-OCT-2000; 2000US-237173P.
 PR 02-OCT-2000; 2000US-237278P.
 PR 02-OCT-2000; 2000US-237294P.
 PR 02-OCT-2000; 2000US-237295P.
 PR 02-OCT-2000; 2000US-237316P.
 PR 02-OCT-2000; 2000US-237425P.
 PR 03-OCT-2000; 2000US-237598P.
 PR 03-OCT-2000; 2000US-237604P.
 PR 03-OCT-2000; 2000US-237606P.
 PR 03-OCT-2000; 2000US-237608P.
 PR 01-NOV-2000; 2000US-244867P.
 PR 01-NOV-2000; 2000US-245084P.

(AVAL-) AVALON PHARM.

XX PA
 XX PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 PI Soppet DR, Weaver Z;
 XX DR WPI; 2002-188264/24.

XX Screening for anti-neoplastic agent involves exposing cells to a
 PT chemical agent to be tested for anti-neoplastic activity, and
 PT determining a change in expression of a gene of a signature gene set -

PS Claim 1; SEQ ID 3524; 44pp; English.

XX The present invention describes a method (M1) for screening for an
 CC anti-neoplastic agent. The method involves exposing cells to a chemical
 CC agent to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening
 CC an anti-neoplastic agent, and can be used for producing a product which
 CC is the data collected with respect to the anti-neoplastic agent as a
 CC result of M1, and the data is sufficient to convey the chemical
 CC structure and/or properties of the agent. M1 can be used in the
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer.
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilms' tumour.

XX Sequence 3184 BP; 921 A; 681 C; 657 G; 925 T; 0 other;

Alignment Scores:

Pred. No.:	2.69e-57	Length:	3184
Score:	596.50	Matches:	117
Percent Similarity:	73.40%	Conservative:	21
Best Local Similarity:	62.23%	Mismatches:	41
Query Match:	56.92%	Indels:	9
DB:	24	Gaps:	2

US-09-782-953-3 (1-198) x ABL65187 (1-3184)

QY 11 SerAlaThrIleAlaCysHisLeuAspProArgValPheValAspGlyLeuCysArgAla 30
 Db 220 TCACCTCTGTCCTGCTGGTGGATGTCGAGCTCTTACCAGTACAGGAGGTTAAGGAA 279
 QY 31 LysPheGluSerLeuPheArgThrTyrAspLysAspThrThrPheGlnTyrPheLysSer 50
 Db 280 AAATTTGGGGGACTGTTTCGGACTTATGATGACTGTGTGAGCTTCAGCTATTAAAGAGT 339
 QY 51 PheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAspAlaArgLeuArg 70
 Db 340 TTCAGAGCTGTCGTATAAATCTTCAGCAATCTTAAATCTGCAGCCGCGAGCTAGGATAGAG 399
 QY 71 LeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyrPheAlaGlnThrLeuHis 90
 Db 400 CTTCAATGAACCCCAATTCAGAGGGGAAAAATTAAGCTCTTACTTTGCACAGGTTTCAGACT 459

QY 91 IleGlySerSer-----HisLeuAlaProProAsnProAsnProAspLysGlnPheLeu 106
 Db 460 CCAGAGACAGATGGAGACAAACTGCACCTTGGCTTCACCCCGAGCTGCCAAACAGTTTCTC 519
 QY 107 IleSerProProAlaSerProProValGlyTrpLysGlnValGluAspAlaThrProVal 126
 Db 520 ATCTCGCCCTTCTCTCCACCTGTTAGCTGGCAGCCCAACGATGCCACGCCAGTC 579
 QY 127 IleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyProGlyGlyLysTyrGluLeu 146
 Db 590 CTCRAACTATGACCTCTCTATGCTGTGCCAAACTAGCACCAGGAGAGAGTATGAGCTC 639
 QY 147 HisAlaAlaThrAspProThrProSerValValHisValCysGluSerAspGlnGlu 166
 Db 640 CATGCAAGGACTGAGTCCACCCCAAGTGTGTCGTCACGTGTGCGACAGTGACATAGAG 699
 QY 167 AsnGluGluGluGluGluMetGluArgMetLysArgProLysProLysIleIleGln 186
 Db 700 GAAGAAGAGAGCCCAAG-----ACTTCCCAAGCCAAAAATCATCATCCA 744
 QY 187 ThrArgArgProGluTyrThrPro 194
 Db 745 ACTCGCGCTCTGCGCTGCCACCC 768

Search completed: December 14, 2002, 20:08:59
 Job time : 145.568 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 14, 2002, 19:56:01 ; Search time 1118.55 Seconds
(without alignments)
2866.843 Million cell updates/sec

Title: US-09-782-953-3
Perfect score: 1048
Sequence: 1 MEEVDLQPLSATIACHLDP.....RPKPKIIQTRPEYTPPIHLS 198

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delopt 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DRV=xlh
-Q/cgr2_1/USPTO_spool/US09782953/runat_11122002_114429_17353/app_query.fasta_1.1173
-DB=EST -QPMI=fastap -SUFFIX=rest -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09782953 @CNC 1.1 1583 @runat_11122002_114429_17353 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_estc:*
9: gb_est1:*
10: gb_est2:*
11: gb_est3:*
12: gb_est4:*
13: gb_est5:*
14: gb_est6:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_Other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1002.5	95.7	828	9	AL536447
2	1002.5	95.7	890	9	AL536447
3	989.5	94.4	867	14	BQ427531
4	974.5	93.0	937	12	BE795722
5	963.5	91.9	956	9	AL556803
6	956.5	91.3	923	9	AL551657
7	904	86.3	1158	14	BQ895506
8	903	86.2	520	14	BQ749142
9	885	84.4	1507	11	AK010696
10	872.5	83.3	837	9	AU124628
11	872.5	83.3	931	9	AL543576
12	872.5	83.3	939	9	AL546617
13	872.5	83.3	1041	13	BM450020
14	865.5	82.6	946	14	BQ278576
15	863.5	82.4	718	9	AU131040
16	859.5	82.0	753	13	BI148584
17	859.5	82.0	885	9	AL544755
18	847.5	80.9	1014	9	AL559594
19	845.5	80.7	906	9	AL544313
20	843.5	80.5	662	13	BI464521
21	842.5	80.4	875	12	EG574693
22	841.5	80.3	939	9	AL554686
23	834.5	79.6	1078	13	BM541636
24	831.5	79.3	740	13	BI463566
25	827.5	79.0	950	12	BF782633
26	820	78.2	710	10	BB617325
27	812.5	77.5	711	13	EG915568
28	812	77.5	769	13	BI767955
29	811.5	77.4	811	12	EG570239
30	809.5	77.2	665	13	BI219142
31	801	76.4	980	12	BQ296537
32	797.5	76.1	905	14	BQ732056
33	790	75.4	827	13	BI102432
34	789.5	75.3	710	12	EG475986
35	784.5	74.9	689	9	AL652662
36	782	74.6	727	13	BI327875
37	776.5	74.1	642	14	BQ520533
38	769.5	73.4	898	14	BQ733563
39	763.5	72.9	751	9	AL576189
40	759.5	72.5	664	13	BJ061844
41	757.5	72.3	526	10	AW957479
42	749.5	71.5	599	9	AA200984
43	747.5	71.3	627	13	BI859506
44	731.5	69.8	1047	12	BQ287042
45	720.5	68.8	599	13	BJ064020

ALIGNMENTS

RESULT 1
AL536447

LOCUS

DEFINITION

AL536447 LTI FL013 FBRn1 Homo sapiens cDNA clone CS0DF015YF21 5

prime, mRNA sequence.

ACCESSION

AL536447.1

VERSION

EST

KEYWORDS

EST

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 828)

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization

TITLE


```

Db 445 ATGAAGTTATATTTCTCAGACCTTACATAGAGAGCTCACCTGGCTCCGCCCAAT 504
Qy 101 ProAspLysGlnPheLeuIleSerProProAlaSerProProValGlyTrpLysGlnVal 120
Db 505 CCAGACAAGCAGTTTCTGATCTCTCCCTCCGCTCTCCGCGAGTGGGATGGAAACAAGTG 564
Qy 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro 140
Db 565 GAAGATGCGACCCAGCTCATAACTATGATCTCTTATATGCCATCTCCCAAGCTGGGGCCA 624
Qy 141 GlyLysLysTyrGlnLeuHisAlaAlaThrAspProThrProSerValValHisVal 160
Db 625 GGGGAAAAGTATGAATTCAGCGAGGAGTGCACCACTCCCGAGCGTGGTGCATGTA 684
Qy 161 CysGluSerAspGlnGluAsnGluGluGluGluMetGluArgMetLysArgPro 180
Db 685 TGTGAGAGTGATCAAGAG---AAGGAGGAGAGAGAGAAATGGAAAGATGAGGAGACCT 741
Qy 181 LysProLysIleGlnThrArgProGluTyrThrProIleHisLeuSer 198
Db 742 AAGCAAAAATATTCAGACCGAGGCGGAGTACAGCGCGATCCACTCAGC 795

RESULT 3
BQ427531 867 bp mRNA linear EST 24-MAY-2002
LOCUS AGENCOURT_7914139 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6156873
DEFINITION 5', mRNA sequence.
ACCESSION BQ427531
VERSION BQ427531.1 GI:21166607
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M13501 row: n column: 10
High quality sequence stop: 626.
FEATURES
source
1..867
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6156873"
/clone_lib="NIH_MGC_71"
/tissue_type="leiomysarcoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: uterus; Vector: pCMV-SPORT6; Site 1: Not I;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2.1 kb."
BASE COUNT 204 a 241 c 267 g 155 t
ORIGIN
Alignment Scores: 3.3e-101 Length: 867
Pred. No.: 989.50 Matches: 191
Score: 97.49% Conservative: 3
Percent Similarity: 95.98% Mismatches: 3
Best Local Similarity: 94.42% Indels: 2
Query Match: 14 Gaps: 1
DB:
US-09-782-953-3 (1-198) x BQ427531 (1-867)

```

```

Qy 1 MetGluGluValAspLeuGlnAspLeuProSerAlaThrIleAlaCysHisLeuAspPro 20
Db 155 ATGAGAGAGGTGGACCTGCGAGGACCTGCCAGCGCCACCATCGCTGTACCTGGACCCG 214
Qy 21 ArgValPheValAspGlyLeuCysArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40
Db 215 CGCGTGTTCGTGGAGCGCTGTGCGGGCCAAATTTAGTCCCTCTTTAGGACGTATGAC 274
Qy 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60
Db 275 AAGACATACACCTTTTCACTATTTTAAGAGCTTCAACAGAGTCAAGATAAATTCAGCAAC 334
Qy 61 ProLeuSerAlaAlaAspAlaArgLeuLeuHisLysThrGluPheLeuGlyLysGlu 80
Db 335 CCCTTCTCCGCGAGCAGATGCCAGGCTCCAGCTGCATAAGACTGAGTTCTTGGGAAAGGAA 394
Qy 81 MetLysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProProAsn 100
Db 395 ATGAAGTTATATTTTGTCTCAGACCTTACATAGAGAGTCCACACTGGCTCCGCCAAAT 454
Qy 101 ProAspLysGlnPheLeuIleSerProProAlaSerProProValGlyTrpLysGlnVal 120
Db 455 CAGACACAGCAGTTTCTGATCTCCCTCCGCTCTCCGCGAGTGGATGAAACAAGTG 514
Qy 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro 140
Db 515 GAAGATGCGACCCAGCTCATAACTATGATCTCTTATATGCCATCTCCCAAGCTGGGGCCA 574
Qy 141 GlyLysLysTyrGlnLeuHisAlaAlaThrAspProThrProSerValValHisVal 160
Db 575 GGGGAAAAGTATGAATTCAGCGAGGAGTGCACCACTCCCGAGCGTGGTGCATGTA 634
Qy 161 CysGluSerAspGlnGluAsnGluGluGluGluMetGluArgMetLysArgPro 180
Db 635 TGTGAGAGTGATCAAGAG---AAGGAGGAGAGAGAGAAATGGAAAGATGAGGAGACCT 691
Qy 181 LysProLysIleGlnThrArgProGluTyrThrProIleHisLeuSer 198
Db 692 AAGCAAAAATATTCAGACCGAGGCGGAGTACAGCGCGATCCACTCAGC 746

RESULT 4
BQ795722 937 bp mRNA linear EST 20-SEP-2000
LOCUS 601590730F1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:3944959 5',
DEFINITION mRNA sequence.
ACCESSION BQ795722
VERSION BQ795722.1 GI:10216920
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: L14M802 row: k column: 08
High quality sequence stop: 812.
FEATURES
source
1..937
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3944959"
/clone_lib="NIH_MGC_7"

```

```

/tissue="small cell carcinoma"
/cell_line="MGC3"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: lung; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAAGAG(G). Size selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

BASE COUNT 215 a 244 c 286 g 192 t

ORIGIN

Alignment Scores:

Pred. No.:	1,76e-99	Length:	937
Score:	974.50	Matches:	190
Percent Similarity:	97.47%	Conservative:	3
Best Local Similarity:	95.96%	Mismatches:	4
Query Match:	92.99%	Indels:	2
DB:	12	Gaps:	1

US-09-782-953-3 (1-198) x BE795722 (1-937)

QY 1 MetGluGluValAspLeuGlnAspLeuProSerAlaThrIleAlaCysHisLeuAspPro 20

DB 37 ATGAGAGAGGTGACCTCGACGACCTGCGCAGCCCACTCCCTGTCACCTGACCCG 96

QY 21 ArgValPheValAspGlyLeuCyArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40

DB 97 CGCGTGTCTCGACGCGCTGCTCCGGCCAAATTGAGTCCCTCTTTAGAGATGAC 156

QY 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60

DB 157 AAGGACATCCCTTTAGATTTTAAAGCTTAAACGAGTCAAGATTAACCTTCACCAAC 216

QY 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheGluGlyLysGlu 80

DB 217 CCTCTCTCCGACGACATGCGCTCCAGCTCCATTAACATGATGTTTGGGAAAGGA 276

QY 81 MetLysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProProAsn 100

DB 277 ATGAAGTTATATTCTCTCAGACCTTACATAGAACTCAACCTGCTCCGCAACT 336

QY 101 ProAspLysGlnPheLeuLysSerProAlaSerProProValGlyTyrIleGlnVal 120

DB 337 CCAGAGAACAGTTTGTATCTCCCTCCGCTCTCCGCGAGTGGATGGAAACAAGTG 396

QY 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro 140

DB 397 GAGATGCGACCCCAAGTCATTAACCTATGATCTTATATGTCATCTCCAAAGCTGGGGCCA 456

QY 141 GlyLysLysTyrGluLeuHisAlaIleThrAspProThrProSerValValHisVal 160

DB 457 GGGAA-AAGTTATGAAATGACGACGACGACTGACACCACTCCAGAGTGGTGTTCATGTA 515

QY 161 CysGluSerAspGlnGluAsnGluGluGluGluGluMetGluArgMetLysArgPro 180

DB 516 TGTGAGAGTGCACAGAG---AAGAGAGAAAGAAAGAAATGGAAAGATGAGAGACT 572

QY 181 LysPhePheValIleIleGlnThrArgArgProGluTyrThrProIleHisLeuSer 198

DB 573 AAGCCAAAAATATATCCAGACCGAGAGCGGAGTACACGCGATCCACTCAGC 626

RESULT 5

AL556803

LOCUS AL556803 956 bp mRNA linear EST 16-FEB-2001

DEFINITION AL556803 LTI_NFL006.PL2 Homo sapiens cDNA clone CSDBK012YD10 5

ACCESSION AL556803

VERSION AL556803.1 GI:12899815

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 956)

AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished (2001)

COMMENT Contact: Genoscope

Genoscope - Centre National de Sequencage

Bp 191 91006 EVRI cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES

source

1..956

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="CSDBK012YD10"

/clone_1db="LTI_NFL006.PL2"

/tissue_type="placenta"

/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by a division of invitrogen 9800 Medical Center Drive, Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@liferetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 214 a 273 c 301 g 158 t 10 others

ORIGIN

Alignment Scores:

Pred. No.:	3.12e-98	Length:	956
Score:	963.50	Matches:	187
Percent Similarity:	95.96%	Conservative:	3
Best Local Similarity:	94.44%	Mismatches:	7
Query Match:	91.94%	Indels:	1
DB:	9	Gaps:	1

US-09-782-953-3 (1-198) x AL556803 (1-956)

QY 1 MetGluGluValAspLeuGlnAspLeuProSerAlaThrIleAlaCysHisLeuAspPro 20

DB 289 ATGAGAGAGGTGACCTCGACGACCTGCGCAGCCCACTCCCTGTCACCTGACCCG 348

QY 21 ArgValPheValAspGlyLeuCyArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40

DB 349 CGCGTGTCTCGACGCGCTGCTCCGGCCAAATTGAGTCCCTCTTTAGAGATGAC 408

QY 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60

DB 409 AAGGACATCCCTTTAGATTTTAAAGCTTAAACGAGTCAAGATTAACCTTCACCAAC 468

QY 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheGluGlyLysGlu 80

DB 469 CCTCTCTCCGACGACATGCGCTCCAGCTCCATTAAGACTGAGATTTTGGGAAAGGA 528

QY 81 MetLysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProProAsn 100

DB 529 ATGAAGTTATATTCTCTCAGACCTTACACATAGAACTCAACCTGCTCCGCAACT 588

QY 101 ProAspLysGlnPheLeuLysSerProAlaSerProProValGlyTyrIleGlnVal 120

DB 589 CCAGACAAAGAGTTTGTATCTCCCTCCGCTCTCCGCGAGTGGGAGGAAACAAGTG 648

QY 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro 140

DB 649 GAGATGCGACCCCAAGTCAATAGATCTCTTATATGCCATCTCCAAAGCTGGGGGCA 708

QY 141 GlyLysLysTyrGluLeuHisAlaIleThrAspProThrProSerValValHisVal 160

DB 709 GGGGAAAGATGAATTTGACGACGAGTACACCACTCCCAAGCTGTGTTCATGTA 768

QY 161 CysGluSerAspGlnGluAsnGluGluGluGluMetGluArgMetLysArgPro 180
 DB 769 TGTGAGAGTGATCNA---RARAAGAGGAGAGAGAGGAATCGAAGATGGAGAGCT 825
 QY 181 LysProLysLeuLeuGlnThrArgArgProGluGluTyrThrProLeuHisLeuSer 198
 DB 826 AAGCCAAATATTATCCARACAGGAGCGGAGTACACGCCGATCCACTCAGC 879

RESULT 6
 AL5511657 AL5511657 923 bp mRNA linear EST 16-FEB-2001
 LOCUS AL5511657 LTI_NFL006.PL2 Homo sapiens cDNA clone CS0D1063YH22 5
 DEFINITION prime, mRNA sequence.
 ACCESSION AL5511657
 VERSION AL5511657.1 GI:12089816
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 923)
 AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
 Location/Qualifiers
 1..923
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS0D1063YH22"
 /clone_lib="LTI_NFL006.PL2"
 /tissue_type="placenta"
 /notes="vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 210 a 257 c 295 g 155 t 6 others
 ORIGIN

Alignment Scores:
 Pred. No.: 1.84e-97 Length: 923
 Score: 956.50 Matches: 188
 Percent Similarity: 96.46% Conservativity: 3
 Best Local Similarity: 94.95% Mismatches: 6
 Query Match: 91.27% Indels: 2
 DB: 9 Gaps: 1

US-09-782-953-3 (1-198) x AL5511657 (1-923)

QY 1 MetGluValAspGlnAspLeuProSerAlaThrIleAlaCysHisLeuAspPro 20
 DB 257 ATGGAGAGGTGGACCTGAGAGACCTGCCAGCGCCACCATCGCTGTCACTGGACCCG 316
 QY 21 ArgValPheValAspGlyLeuCysArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40
 DB 317 CGCGTGTCTGTGGAGCGCTGTGCCGGGCAAAATTTAGTCCCTCTTTAGGACGTATGAM 376
 QY 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60
 DB 377 AAGGACATCACCTTTTCAGTATTTTAAAGACTTCAACAGAGTCAAGATAAATTCAGCAAC 436
 QY 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
 DB 437 CCCTTCTCCGACAGATGCCAGGCTCCAGCTGCATTAARACTGAGTTCTCTGGGAAGGNA 496

QY 81 MetLysLeuTyrPheAlaGlnThrLeuHisLleGlySerSerHisLeuAlaProAsn 100
 DB 497 ATGAAGTTATATTTTCTCAGACCTTACACATAGGAAGCTCACCTGGCTCCGCCAAAT 556
 QY 101 ProAspLysGlnPheLeuLeuSerProProAlaSerProProValGlyTyrLysGlnVal 120
 DB 557 CCGACACAGCAGTTTCTGATCTCCCTCCCGCTCTCCCGCAGTGGGATGGAAACAGTG 616
 QY 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaLleSerLysLeuGlyPro 140
 DB 617 GAAGTCCGACCCCACTCATAACTATGATCTCTATATGCCATCTCCAAGCTGGGGCCA 676
 QY 141 GlyGluLysTyrGluLeuHisAlaIaThrAspProThrProSerValValHisVal 160
 DB 677 GCGGAAAGATGAAATTCACGCGAGCTGCACCACTCCCGAGTGGTGGTCCATGTA 736
 QY 161 CysGluSerAspGlnGluAenGluGluGluGluMetGluArgMetLysArgPro 180
 DB 737 TGTGAGAGTGATCAAGAG---AAGGAGAGAGAGGAATGGAAAGATGAGGAGACT- 792
 QY 181 LysProLysLleGlnThrArgArgProGluTyrThrProLeuHisLeuSer 198
 DB 793 AAGCCAAATATTATCCAGACGAGGCGGAGTACACGCCGATCCACTCAGC 846

RESULT 7
 BQ895506 BQ895506 1158 bp mRNA linear EST 16-AUG-2002
 LOCUS BQ895506 AGENCOURT 8762889 NIH_MGC_130 Mus musculus cDNA clone IMAGE:6329174
 DEFINITION 5', mRNA sequence.
 ACCESSION BQ895506
 VERSION BQ895506.1 GI:22287520
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 1158)
 AUTHORS NIH-MGC http://mgs.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman, Ph.D.
 cDNA Library Preparation: ResGen, Invitrogen Corp
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM13781 row: i column: 15
 High quality sequence start: 83
 High quality sequence stop: 577.
 Location/Qualifiers
 1..1158
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="IMAGE:6329174"
 /lab_host="NIH MGC_130"
 /note="Organ: oocytes; Vector: pCMV-SPORT6.1.ccdB; Site 1: EcoRV; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.95 kb. Constructed by ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

BASE COUNT 289 a 354 c 288 g 226 t 1 others
 ORIGIN

Alignment Scores:
 Pred. No.: 1.95e-91 Length: 1158
 Score: 904.00 Matches: 173
 Percent Similarity: 94.68% Conservativity: 5

Best Local Similarity: 92.02% Mismatches: 10
 Query Match: 86.26% Indels: 0
 DB: 14 Gaps: 0

US-09-782-953-3 (1-198) x BQ895506 (1-1158)

QY 11 SerAlaThrTrpLeuAlaCysHisLeuAspProArgValPheValAspGlyLeuCysArgAla 30
 DB 139 AGCTCCGATTTGGCTTGTGGCAACGATGATGCTTCCAGCCAAAGTGAAGACCGGCC 198
 QY 31 LysPheGluSerLeuPheArgThrTyrAspLysAspThrPheGlnTyrPheLysSer 50
 DB 199 AATTTGATATCTCTTCCAGAACATATGACAAAGACACCCCTTCCAGTATTTAAAGAC 258
 QY 51 PheLysArgValArgIleAsnPheSerAsnProLeuSerAlaIleAspAlaArgLeuArg 70
 DB 259 TTCAACCGTCCCGATTAACCTCAGCAACCCCTTATCTGAGCGGATGCGAGCTCCGG 318
 QY 71 LeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyrPheAlaGlnThrLeuHis 90
 DB 319 CTGCACAAAGCCGAGTTCCTGGGAGAAATGAAGTTGTAATTTGCTCAGACTTTACAC 378
 QY 91 IleGlySerSerHisLeuAlaProProAsnProAspLysGlnPheLeuIleSerProPro 110
 DB 379 ATAGGAAGTTCACACCTGGCTCCGCCCAATCCGACAAAGTTCCTCATCTCCCTCCG 438
 QY 111 AlaSerProProValGlyTyrLysGlnValGluAspAlaThrProValIleAsnTyrAsp 130
 DB 439 GCTCTCTCTCTCTCTGCGTGGAAACAGTAGAAGATGCCACCCCTCATTAATTACGAT 498
 QY 131 LeuLeuTyrAlaIleSerLysLeuGlyProGlyGluLysTyrGluLeuHisAlaAlaThr 150
 DB 499 CTTTATATGCCATCTCCAGAGCTGGGCCGAGAGAAATGATACCTGATGAGCGACA 558
 QY 151 AspProThrProSerValValHisValCysGluSerAspGlnGluAsnGluGlu 170
 DB 559 GACACCACTCCAGTGTGTGTGTCCAGTGTGTGAGAGTGAACCAAGATGAGAGAGAA 618
 QY 171 GluGluGluMetGluArgMetLysArgProLysProLysIleIleGlnThrArgArgPro 190
 DB 619 CAGGAAGAGTGGAGAAATGAAGAGACCAACCCCAATCATCCAGACAGAGACCG 678
 QY 191 GluTyrThrProIleHisLysSer 198
 DB 679 GAGTACACACCCATCCACCTCAGC 702
 RESULT 8
 LOCUS BQ749142 520 bp mRNA linear EST 17-JUN-2002
 DEFINITION UI-M-PD0-BYD-C-04-0-UI.r1 NIH BMAP_PD0 Mus musculus cDNA clone
 IMAGE:5716539 5', mRNA sequence.
 ACCESSION BQ749142
 VERSION BQ749142.1 GI:21895929
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-rcmail.nih.gov
 Tissue Procurement: Dr. James Lahn, University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LINL at:
 http://image.llnl.gov
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

FEATURES
 source
 Seq primer: pYX-5.
 Location/Qualifiers
 1..520

/organism="Mus musculus"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="MIMAG:5716539"
 /clone_1db="NIH_BMAP_PD0"
 /tissue_type="whole brain"
 /dev_stage="embryo 12.5 dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /note="Organ: brain; Vector: pYX-Aec; Site 1: EcoR I;
 Site 2: Not I. The library was constructed according to
 Bonaldi, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured mRNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with an
 oligo-dT primer containing a Not I site. Double stranded
 cDNA was size selected according to mRNA size fraction,
 ligated with EcoR I adaptor, digested with Not I, and then
 cloned directionally into pYX-Aec vector. The library tag
 sequence located between the Not I site and the polyA tail
 is TAGAGAGCC. This library was created for the
 University of Iowa Mouse Brain Molecular Anatomy Project
 (BMAP): 'Gene discovery in the developing Mouse Nervous
 System', supported by National Institutes of Mental Health
 (NIMH), Hemlin Chin, Ph.D., program coordinator."

BASE COUNT
 ORIGIN

Alignment Scores:

Pred. No.: 1,03e-91 Length: 520
 Score: 903.00 Matches: 171
 Percent Similarity: 99.42% Conservative: 0
 Best Local Similarity: 99.42% Mismatches: 1
 Query Match: 86.16% Indels: 0
 DB: 14 Gaps: 0

US-09-782-953-3 (1-198) x BQ749142 (1-520)

QY 8 AspLeuProSerAlaThrTrpLeuAlaCysHisLeuAspProArgValPheValAspGlyLeu 27
 DB 3 GACCTGCCAGAGCCGACCATGCTGCTGCACCTGACCCGCGCTGTCGAGCGGCTG 62
 QY 28 CysArgAlaLysPheGluSerLeuPheArgThrTyrAspLysAspThrPheGlnTyr 47
 DB 63 TGCCGGCCAAATTTGATCTCTTCCAGAACATATGACAAAGACACCACTTCAGTAT 122
 QY 48 PheLysSerPheLysArgValArgIleAsnPheSerAsnProLeuSerAlaIleAspAla 67
 DB 123 TTTAAGAGCTTCAACGCTGCCGATTAACCTCAGCAACCCCTTATCTCAGCCGATGCC 182
 QY 68 ArgLeuArgLeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyrPheAlaGln 87
 DB 183 AGGCTGGCGCTGCACAAAGCCAGATTCCTGGGAGAAATGAATGTAATTTGCTCAG 242
 QY 88 ThrLeuHisIleGlySerSerHisLeuAlaProProAsnProAspLysGlnPheLeuIle 107
 DB 243 ACTTACACATAGAGATTCACACCTGCTCCGCCCAATCCGACAAACAGTTCCTCATC 302
 QY 108 SerProAlaSerProProValGlyTyrLysGlnValGluAspAlaThrProValIle 127
 DB 303 TCCCTCCGCGCTCTCTCCGTGGTGGAAACAAAGTAGAAGATGCCACCCCGTCATA 362
 QY 128 AsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyProGlyGluLysTyrGluLeuHis 147
 DB 363 AATTAGATCTTTATATCCATCTCCAGCTGGGCCGAGAGAAATGATGAATGATCAT 422
 QY 148 AlaAlaThrAspProThrProSerValValHisValCysGluSerAspGlnGluAsn 167
 DB 423 GCAGCAGACAGACCACTCCGAGTGTGTGTGTCCAGTGTGTGAGAGTGAACCAAGAGAT 482
 QY 168 GluGluGluGluGluMetGluArgMetLysArg 179

Db	55	GCCAATTGGAAATCCCTCTTCAGAAACATATGACAAGACACCACTTCAGATTTTAAAG	114
Qy	50	SeRPheLYaRGyAlaYArgLIeaRPheSeRPheProLeuSeRAlaAlaSPheAlaArgLeu	69
Db	115	AGCTTAAACGGTGTCCGGATTAACCTTCAGAACCCCTTATCTGCAGCCCAATGCCAGGCTG	174
Qy	70	ArgLeuHLeuYThGtIuphLeuGtLYLVSGLuMeLYLeuYzPheAlaGInThLeu	89
Db	175	CGGCTSCACAAAGACCGAGTTCCTGGGGAAAGAAATGAAGTGTATTTGGTCAGACTTTA	234
Qy	90	HLeuEGLYSeSeRnHLeuAlaPProPoleRPProAspLYSGLInPheLeuIleSeRPro	109
Db	235	CACATAGGAAGTTCACACCTGGCTCGGCCCAATCCGACAAACAGTTCTCATCTCCCT	294
Qy	110	ProAlaSeRProRoValGtYTrPLYSGInValGtAluAspAlaThrProValIleAsnYr	129
Db	295	CCGGCTCTCCCTCCGTTGGCTGGAAACAAGTAGAAGATGCCACCCCTCATTAATTATC	354
Qy	130	AppLeuLeuTYrAlaIleSeRlyLeuGLYProGLYLuYATYrGLuLeuHIsAlaAla	149
Db	355	GATCTTTTATATGCCATCTTCACAGCTGGGGGCCAGAGAAAGATATACATGCATGCAGCG	414
Qy	150	ThaAPProThProSeRAlValValHIsValCYAGLSuSeRPoGInLuAnGInGLu	169
Db	415	ACAGACACCACTCCCAAGTGTGTGTGTCTCCAGTGTGTGAAGTGAACCAAGAAATGAGAG	474
Qy	170	GLuGLuGLuGLuMeGLuARGMeCYLuarGMeCYLuarGProLYsProLYsIleIleGInThArg	189
Db	475	GAAAGAGAAAGATGGAGAGAAATGAAGAGACCAACCCAAATATCTCAGACACGGAGA	534
Qy	190	ProGLuYrThProIleHIsLeuSeR	198
Db	535	CCGAGATCACACCCATCCCACTTCAGC	561

LOCUS	837 bp	mRNA	linear	EST 01-NOV-2002
DEFINITION	AU124628	NT2RM4	Homo sapiens CDNA clone NT2RM4000307 5', mRNA	
ACCESSION	AU124628			
VERSION	AU124628.1	GI:10945344		
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
AUTHORS	1 (bases 1 to 837) Ota, T., Wakamatsu, A., Ozawa, M., Ishii, S., Saito, K., Yamamoto, J., Nakamura, Y., Nishikawa, T., Nagai, T., Suzuki, Y., Sugano, S. and Iisaga, T.			
TITLE	Hrt human cDNA project (Ota, T., Wakamatsu, A., Ozawa, M., Ishii, S., Saito, K., Yamamoto, J., Nakamura, Y., Nishikawa, T., Nagai, T., Suzuki, Y., Sugano, S., Iisaga, T.)			
JOURNAL	Unpublished (2000)			
COMMENT	Contact: Takao Iisaga Genomics Laboratory Helix Research Institute 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan Tel.: 81-438-52-3975 Fax: 81-438-52-3986 Email: genomics@hri.co.jp HRI human cDNA project: 5'- & 3'-end one pass sequencing: Helix Research Institute, cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.			

FEATURES

SOURCE

```
Location/Qualifiers
1. . 837
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NT2RM4000307"
/clone_lib="NT2RM4"
/cell_type="teratocarcinoma"
/cell_line="NT2"
```

```

/notes="Vector: pME18SFJ3; mRNA from uninduced NT2 neuronal
precursor cells"
BASE COUNT      241 a      201 c      207 g      186 t      2 others
ORIGIN

```

Alignment Scores:	
Pred. No.:	4,78e-8
Score:	872.50
Percent Similarity:	93.0%
Best Local Similarity:	89.36
Query Match:	83.25
DB:	9
	1
Gaps:	1

US-09-782-953-3 (1-198) X AU124628 (1-837)

Oy	11	SerIATrrIIeAlACVSHleuAbprProArValPheValAspGlyLeucYArgLa	30
Db	148	AGCTCCCGATTGGCTGTGTGGCAACAGCATATCTTTCACGGAAAGTGAACCAGGCC	207
Oy	31	LysPheGIuSerLeuPheAlqThrTYrAspLysAphTrhPheGIInTyPhelYsser	50
Db	208	AAATTTGATGCTCCTTTTAGAGCGTATGCACAAGACATCACCTTTCAGTATTTTAAGAGC	267
Oy	51	PheLYsArGVAlArGIleAsnPheSerAnPrProLeuSerAlAlAspAlArgLeuArg	70
Db	268	TTCAAACAGACGTACAGATTAATCTTCAGCAACCCCTTCTCCGACGACGATGCCAGGCTCCAG	327
Oy	71	LeuHlSlyThrTGluPheLeuGIlyLysGIuMetLysLeuTYrPheAlagInThrlEuHls	90
Db	328	CTGATATAGACTGTGATTTCTGGAAAGGAAATGAAAGTATATTTTGCTCAACCTTACAC	387
Oy	91	IleGIysSerSerHlSleuAlAlProProAbnPrOApLyrgInPheLeuIlleSerProPro	110
Db	388	ATAGAAGCTCACACCTGTGCTCCGCCCAATCCAGACAGACAGCTTTCGATCTCCCTCCC	447
Oy	111	AlAserProProValGIlyTPDylsGIlnvalGIuAparAlathrProValIlleAsnTYrAsp	130
Db	448	GCCCTCTCCGACAGTGGGATGTGMAAACAGTGGAMATGCGAACCAGTCATAACTATGAT	507
Oy	131	LeuLeuTYrAlAlleSerLYsLeuGIlyProGIlyLysTYrGIuLeuHlAlAlAlthr	150
Db	508	CTCTTATATGTCATCTCCMACCTGGGGCCAGGGAAAAGTATGAATTCATGACAGCGCACT	567
Oy	151	AspProThrProSerValAlvalHlsValCYesGIuSerAspGIuIngluasnGIuGIu	170
Db	568	GACACCCACTCCACCGTGGTGTCTCATGTATGTAGATGATCAAGAG--AAGAGAGAA	624
Oy	171	GIuGIuGIuMetGIuAYsMeLYsArgProLYsAprProLYselleGIInThrAlayArgPro	190
Db	625	GAAAGAGAAATTCGAAAGAAATGAGAGACTTAAAGCAAAAAATTATCCAGACCAGAGAGCGC	684
Oy	191	GIuTYrThrProIlleHlsLeuSer	198
Db	685	GAGTACACGCCGATCCACCTCAGC	708
RESULT 11			
LOCUS	AL543576	931 bp	mRNA linear EST 16-FEB-2001
DEFINITION	AL543576 LIT_NFL006_PU2 Homo sapiens cDNA clone CS01006YK09 5		
ACCESSION	AL543576		
VERSION	AL543576.1		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Butleria; Primates; Catarrhini; Homidae; Homo.			
REFERENCE	Li,W.B., Gruber,C., Jesse,J. and Polyes,D.		
AUTHORS	Full-length cDNA libraries and normalization		
TITLE	Unpublished (2001)		
JOURNAL	Contact: Genome		
COMMENT	Genoscope - Centre National de Sequenage		

BP 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES

Source
1. .931
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSODI029XJ21"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 262 a 217 c 238 g 214 t

ORIGIN

Alignment Scores:
Pred. No.: 5.39e-88 Length: 931
Score: 872.50 Matches: 168
Percent Similarity: 93.09% Conservative: 7
Best Local Similarity: 89.36% Mismatches: 12
Query Match: 83.25% Indels: 1
DB: 9 Gaps: 1

US-09-782-953-3 (1-198) x AL543576 (1-931)

QY 11 SerAlaThrIleAlaCysHisLeuAspProArgValPheValaspGlyLeuCysArgAla 30
Db 80 AGCTCCCTGATTGCTGTGGCAACAGTGATATCTTCAGCAAGAGTGAACAGGGCC 139
QY 31 LysPheGluSerLeuPheArgThrTyrAspLysAspThrPheGlnTyrPheLysSer 50
Db 140 AATTTGAGTCCCTCTTTAGGAGTATGACAGGACATCACCTTTTCAGTATTTTAAGAGC 199
QY 51 PheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAspAlaArgLeuArg 70
Db 200 TTCAACAGCAGTCAGATTAATACTTCAGCAACCCCTTCTCCGACGAGATGCCAGGCTCCAG 259
QY 71 LeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyrPheAlaGlnThrLeuHis 90
Db 260 CTGCATGAAGCTGAGTTCTTGGGAAGAAATGAAGTATATTTTGTCTCAGACCTTACAC 319
QY 91 IleGlySerSerHisLeuAlaProAsnProAspLysGlnPheLeuIleSerProPro 110
Db 320 ATAGGAAGCTCACACCTGGCTCGCCAAATCCACAGAGCAGTTTCTGATCTCCCTCC 379
QY 111 AlaSerProValGlyTyrLysGlnValGluAspAlaThrProValIleAsnTyrAsp 130
Db 380 GCCTCTCCGCGAGTGGATGGAACAAGTGAAGATGCGACCCAGTCATAAATATGAT 439
QY 131 LeuLeuTyrAlaIleSerLysLeuGlyProGlyLysLysTyrGluLeuHisAlaAlaThr 150
Db 440 CTCTTATATGCACTCTCCAGCTGGGCGAGGAAAGTGAATGAATGTCAGCGAGCACT 499
QY 151 AspProThrProSerValValHisValCysGluSerAspGlnGluAsnGluGlu 170
Db 500 GACACCACTCCAGCGTGGTCCATGATGTCAGAGTGTATCAAGAG---AAGGAGGAA 556
QY 171 GluGluMetGluArgMetLysArgProLysProLysIleIleGlnThrArgArgPro 190
Db 557 GAAGAGGAAATGGAAGAAATGAGAGACCTAAGCCAAAAATATCCAGCAGCAGGAGCCA 616
QY 191 GluTyrThrProIleHisLeuSer 198
Db 617 GAGTACAGCGCGGATCCACCTCAGC 640

RESULT 12

AL546617
LOCUS

DEFINITION AL546617 LTI_NFL006_PL2 Homo sapiens cDNA clone CSODI029YJ21 5 prime, mRNA sequence.

ACCESSION AL546617

VERSION AL546617.1 GI:12879906

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 939)

AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished (2001)

COMMENT Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES

Location/Qualifiers

1. .939

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="CSODI029XJ21"

/clone_lib="LTI_NFL006_PL2"

/tissue_type="placenta"

/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 263 a 219 c 238 g 218 t

ORIGIN

Alignment Scores:

Pred. No.: 5.44e-88 Length: 939

Score: 872.50 Matches: 168

Percent Similarity: 93.09% Conservative: 7

Best Local Similarity: 89.36% Mismatches: 12

Query Match: 83.25% Indels: 1

DB: 9 Gaps: 1

US-09-782-953-3 (1-198) x AL546617 (1-939)

QY 11 SerAlaThrIleAlaCysHisLeuAspProArgValPheValaspGlyLeuCysArgAla 30
Db 83 AGCTCCCTGATTGCTGTGGCAACAGTGATATCTTCAGCAAGAGTGAACAGGGCC 142
QY 31 LysPheGluSerLeuPheArgThrTyrAspLysAspThrPheGlnTyrPheLysSer 50
Db 143 AATTTGAGTCCCTCTTTAGGAGTATGACAGGACATCACCTTTTCAGTATTTTAAGAGC 202
QY 51 PheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAspAlaArgLeuArg 70
Db 203 TTCAACAGCAGTCAGATTAATACTTCAGCAACCCCTTCTCCGACGAGATGCCAGGCTCCAG 262
QY 71 LeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyrPheAlaGlnThrLeuHis 90
Db 263 CTGCATGAAGCTGAGTTCTTGGGAAGAAATGAAGTATATTTTGTCTCAGACCTTACAC 322
QY 91 IleGlySerSerHisLeuAlaProAsnProAspLysGlnPheLeuIleSerProPro 110
Db 323 ATAGGAAGCTCACACCTGGCTCGCCAAATCCACAGCAGTGTTCGATCTCCCTCC 382
QY 111 AlaSerProValGlyTyrLysGlnValGluAspAlaThrProValIleAsnTyrAsp 130
Db 383 GCCTCTCCGCGAGTGGATGGAACAAGTGAAGATGCGACCCCGAGTCATAAATATGAT 442

```

Qy 131 LeuLeuTyraIaIleSerLyLeuGlyProGlyGluLyTyGluLeuHlaIaIaThr 150
Db 443 CTCCTATATGCGCATCTCCAAAGCTGGGGCCAGGGAAGATATGATTCACGACGACCT 502
Qy 151 AppProThProSerValValHlaIaIaValCyAGluSerAspGlnGluangluGlu 170
Db 503 GACACCACTCCCGCTGGTGGTCATGTATGTGAAGATGATCAAGAG--AAGGAGGA 559
Qy 171 GluGluGluMetGluArgMetLyAspProLyAspProLyIleIleGlnThrArgArgPro 190
Db 560 GAAGAGGAATGGAAGATGAGAGAGCTTAAGCCAAAATTATTCAGACGAGGAGCGG 619
Qy 191 GluTyThrProIleHlaIeUser 198
Db 620 GAGTACACGCCGATCCACCTCAGC 643

RESULT 13
BM450020 1041 bp mRNA linear EST 05-FEB-2002
LOCUS BM450020
DEFINITION AGENCOURT_6393352 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5528025
5', mRNA sequence.
ACCESSION BM450020
VERSION BM450020.1 GI:18499060
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1041)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/BTP
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM12204 row: d column: 10
High quality sequence stop: 643.
FEATURES
source
1..1041
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5528025"
/clone_lib="NIH_MGC_72"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: skin; Vector: PCMV-SPORT6; Site:1: Nci;
Site:2: Sall; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."
BASE COUNT 282 a 248 c 274 g 236 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 6.1e-88 Length: 1041
Score: 872.50 Matches: 168
Percent Similarity: 93.09% Conservative: 7
Best Local Similarity: 89.36% Mismatches: 12
Query Match: 83.25% Indels: 1
DB: 13 Gaps: 1
US-09-782-953-3 (1-198) x BM450020 (1-1041)
Qy 11 SerAlaThrTAlaIaCyHlaIeUserProArgValPheValAspGlyLeuCyArgAla 30
Db 114 AGCTCCCGATGGCTGGTGGTGGCAAGATGATATCTTCAGCAAGTGAACGAGGCC 173
Qy 31 LybPheGluSerLeuPheArgThrTyraAspLyAspThrThrPheGlnTyrPheYasSer 50

```

```

Db 174 AAATTGAGTCCCTCTTAGAGCGATGACAAAGACATCACCTTCAGTATTTTAAGAC 233
Qy 51 PheLyAspArgValArgIleAsnPheSerAsnProLeuSerAlaIaAspAlaArgLeuArg 70
Db 234 TTCAAACGAGTCAGATAACTTCAGCAACCCCTTCGCGAGCAGATGCGAGGCTCCAG 293
Qy 71 LeuHlaIeTyThrGluPheLeuGlyLyAspGluMetLyLeuTyThrPheAlaGlnThrLeuHla 90
Db 294 CCGCATTAAGACTGAGATTTCTGGGAAGAAATGAAGTTATTTTGTCTCAACCTTACAC 353
Qy 91 IleGlySerSerHlaIeAlaProProAsnProAspLyGlnPheLeuIleSerProPro 110
Db 354 ATAGAGAGCTCACACTGCTCCGCAATCCAGACAGAGTTTCTGATCTCCCTCC 413
Qy 111 AlaserProProValGlyTyThrLyGlnValGluAspAlaThrProValIleAsnTyraAsp 130
Db 414 GCTCTCCGCGAGTGGATGGAAGAAACAGTGAAGATGGACCCAGTCATTAACCTATGAT 473
Qy 131 LeuLeuTyraIaIleSerLyLeuGlyProGlyGluLyTyGluLeuHlaIaIaThr 150
Db 474 CTCCTATATGCGCATCTCCAAAGCTGGGGCCAGGGAAGATATGATTCACGACGACCT 533
Qy 151 AppProThProSerValValHlaIaIaValCyAGluSerAspGlnGluangluGlu 170
Db 534 GACACCACTCCCGCTGGTGGTCATGTATGTGAAGATGATCAAGAG--AAGGAGGA 590
Qy 171 GluGluGluMetGluArgMetLyAspProLyAspProLyIleIleGlnThrArgArgPro 190
Db 591 GAAGAGGAATGGAAGATGAGAGAGCTTAAGCCAAAATTATTCAGACGAGGAGCGG 650
Qy 191 GluTyThrProIleHlaIeUser 198
Db 651 GAGTACACGCCGATCCACCTCAGC 674

RESULT 14
B0278576 946 bp mRNA linear EST 07-MAY-2002
LOCUS B0278576
DEFINITION AGENCOURT_6773220 NIH_MGC_128 Homo sapiens cDNA clone IMAGE:5811711
5', mRNA sequence.
ACCESSION B0278576
VERSION B0278576.1 GI:20488772
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 946)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: NCI
CDNA Library Preparation: Michael Brownstein Laboratory
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM2060 row: p column: 16
High quality sequence stop: 548.
FEATURES
source
1..946
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5811711"
/clone_lib="NIH_MGC_128"
/tissue_type="mixed (pool of 40 RNAs)"
/lab_host="DH10B (T1-phage-resistant)"
/notes="Vector: pDNR-LIB; Site:1: SfiI (ggccatcatggcc);
Site:2: SfiI (ggcgccctcgcc); Double-stranded cDNA was
prepared from a pool of 40 cell line polyA+ RNAs (bladder

```

- 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon - 4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%, kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell - 5.2%, ovary - 4%, pharynx - 2.5%, prostate - 4.3%, salivary gland - 1.3%, and skin - 2.3%. 5' and 3' adaptors were used in cloning as follows:

5'-AAGCAGTGGTATCAACGAGAGTGGCCATTACGCGGG-3' and 5'-ATTCTAGAGCCGAGCGGCCGACATG-TT(30)NN-3'. Full-length enriched library was constructed using the Clontech Creator SMART kit and size-selected to contain the >2 kb size fraction (other fractions present in NIH_MGC_126 and NIH_MGC_127). Library created in the laboratory of T. Usdin, M.D., Ph.D. (NIH, NIH). Note: this is a NIH_MGC Library."

BASE COUNT 264 a 243 c 239 g 200 t
ORIGIN

Alignment Scores:

Pred. No.: 3,378-87 Length: 946
Score: 865.50 Matches: 167
Percent Similarity: 92.02% Conservative: 6
Best Local Similarity: 88.83% Mismatches: 14
Query Match: 82.59% Indels: 1
DB: 14 Gaps: 1

US-09-782-953-3 (1-198) x BQ278576 (1-946)

Qy 11 SerAlaThrIleAlaCysHisLeuAspProArgValPheValAspGlyLeuCysArgAla 30
Db 106 AGCTTCCTGATGCTGTGGCAACAGTGATATCTTCAGCGAAGAGTGAACACAGGGCC 165
Qy 31 LysPheGluSerLeuPheArgThrTyrAspLysAspThrThrPheGlnTyrPheLysSer 50
Db 166 AAATTGAGTCCCTCTTAGACGCTATGACAAAGACATCACCTTTTCAGTATTTTAAAGC 225
Qy 51 PheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAspAlaArgLeuArg 70
Db 226 TTCAACAGAGTCAGATAAACTTCAGCAACCCCTTCTCCGAGCAGATGCCAGGCTCCAG 285
Qy 71 LeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyrPheAlaGlnThrLeuHis 90
Db 286 CTGATAGACTGAGTTCTGGGAAGAAATGAAGTTATATTTTGTCTGACCTTACAC 345
Qy 91 IleGlySerSerHisLeuAlaProAsnProAspLysGlnPheLeuIleSerProPro 110
Db 346 ATAGGAAGCTCACACCTGGCTCGCCAAATCCAGCAAGCAGTTTCTGATCTCCCTCCC 405
Qy 111 AlaSerProProValGlyTyrLysGlnValGluAspAlaThrProValIleAsnTyrAsp 130
Db 406 GCCTCTCCGCGAGTGGATGGAACAGTGAAGATGCCACCCAGTCATAAACTATGAT 465
Qy 131 LeuLeuTyrAlaIleSerLysLeuGlyProGlyGluLysTyrGluLeuHisAlaAlaThr 150
Db 466 CTCATTATATGCAATCTCAAGCTGGGGCCAGGGAAAGATGAATTCAGCGAGGACT 525
Qy 151 AspProThrProSerValValHisValCysGluSerAspGlnGluAsnGluGluGlu 170
Db 526 GACACCACTCCAGCGTGGTCCATGATGATGAGAGTGAAGAAC--GAGGAA 582
Qy 171 GluGluGluMetGluArgMetLysArgProLysProLysIleIleGlnThrArgArgPro 190
Db 583 GAAGAGGAAATGGAAGAATGACGAGACTTAAGCCAAAAATTTATCCCGACCGAAGCCG 642
Qy 191 GluTyrThrProIleHisLeuSer 198
Db 643 GAGTACACGCCGATCCACTCAGC 666

RESULT 15

AUI31040
LOCUS AUI31040 NT2RP3 Homo sapiens cDNA clone NT2RP3001895 5', mRNA
DEFINITION AUI31040 NT2RP3 Homo sapiens cDNA clone NT2RP3001895 5', mRNA
sequence.
ACCESSION AUI31040

VERSION KEYWORDS SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AUI31040.1 GI:10991394

EST.

human.

Homo sapiens

Eukaryota; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 718)

Ota, T., Nishikawa, T., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y.,

Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagai, T., Sugano, S. and

Isogai, T.

HRI human cDNA project

Unpublished (2000)

Contact: Takao Isogai

Genomics Laboratory

Helix Research Institute

1532-3 Yana, Kisarazu, Chiba 292-0812, Japan

Tel: 81-438-52-3975

Fax: 81-438-52-3986

Email: genomics@hri.co.jp

HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix

Research Institute; cDNA library construction: Department of

Virology, Institute of Medical Science, University of Tokyo, and

Helix Research Institute.

Location/Qualifiers

1..718

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="NT2RP3001895"

/clone_lib="NT2RP3"

/cell_type="teratocarcinoma"

/cell_lines="NT2"

/note="Vector: pME18SFL3; mRNA from NT2 neuronal precursor

cells after 2-weeks retinoic acid (RA) induction"

BASE COUNT 217 a 174 c 169 g 155 t 3 others

ORIGIN

Alignment Scores:

Pred. No.: 4,168-87 Length: 718
Score: 863.50 Matches: 166
Percent Similarity: 92.55% Conservative: 8
Best Local Similarity: 88.30% Mismatches: 13
Query Match: 82.40% Indels: 1
DB: 9 Gaps: 1

US-09-782-953-3 (1-198) x AUI31040 (1-718)

Qy 11 SerAlaThrIleAlaCysHisLeuAspProArgValPheValAspGlyLeuCysArgAla 30
Db 132 AGCTTCCTGATGCTGTGGCAACAGTGATATCTTCAGCGAAGAGTGAACACAGGGCC 191
Qy 31 LysPheGluSerLeuPheArgThrTyrAspLysAspThrThrPheGlnTyrPheLysSer 50
Db 192 AAATTGAGTCCCTCTTAGGAGCTATGACAGGACATCACCTTTTCAGTATTTTAAAGC 251
Qy 51 PheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAspAlaArgLeuArg 70
Db 252 TTCAACAGAGTCAGATAAACTTCAGCAACCCCTTCTCCGAGCAGATGCCAGGCTCCAG 311
Qy 71 LeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyrPheAlaGlnThrLeuHis 90
Db 312 CTGATAAGACTGAGTTCTGGGAAGAAATGAAGTTATATTTTGTCTGACACTTACAC 371
Qy 91 IleGlySerSerHisLeuAlaProAsnProAspLysGlnPheLeuIleSerProPro 110
Db 372 ATAGGAAGCTCACACCTGGCTCGCCAAATCCAGCAAGCAGTTTCTGATCTCCCTCCC 431
Qy 111 AlaSerProProValGlyTyrLysGlnValGluAspAlaThrProValIleAsnTyrAsp 130
Db 432 GCCTCTCCGCGAGTGGATGGAACAGTGAAGATGCCACCCAGTCATAAACTATGAT 491
Qy 131 LeuLeuTyrAlaIleSerLysLeuGlyProGlyGluLysTyrGluLeuHisAlaAlaThr 150
Db 492 CTCTTATATGCCATCTCAAGCTGGGGCCAGGGAAAGATGATGAATTGCACGAGGACT 551

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 11:37:40 ; Search time 13.0219 Seconds
(without alignments)
1461.736 Million cell updates/sec

Title: US-09-782-953-3
Perfect score: 1048
Sequence: 1 MEEVDLQPSATIACHLDP.....RPKPKLIQTRRPEYPIHLS 198

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 73:.*
1: Piri:.*
2: Piri2:.*
3: Piri3:.*
4: Piri4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	355	33.9	239	T34305	hypothetical prote
2	139.5	13.3	315	T49641	nebula related pro
3	95.5	9.1	163	S62444	conserved hypothet
4	94	9.0	318	C84651	hypothetical prote
5	92	8.8	992	T46337	hypothetical prote
6	91.5	8.7	1274	D84485	probable retroelem
7	88	8.4	284	B48516	surfactant protein
8	88	8.4	630	S29796	hypothetical prote
9	88	8.4	721	S29795	hypothetical prote
10	86.5	8.3	822	S30198	Na+/H+-exchanging
11	86	8.2	338	S73902	UDP-glucose 4-epim
12	84.5	8.1	719	S51739	transcription repr
13	83	7.9	561	JC2436	5'-nucleotidase (E
14	83	7.9	1173	T31421	C-terminal domain-
15	82	7.8	697	S96752	hypothetical prote
16	82	7.8	1477	T13797	tumor suppressor pr
17	81.5	7.8	528	C69323	conserved hypothet
18	81	7.7	151	T10768	latex allergen Hev
19	80.5	7.7	280	C64471	hypothetical prote
20	80.5	7.7	360	SU2850	protein tpx - Ther
21	80	7.6	681	T17342	hypothetical prote
22	79.5	7.6	797	A70453	glutamate ammonia
23	79.5	7.6	904	I38757	homolog of Drosoph
24	79.5	7.6	926	I38756	homolog of Drosoph
25	79.5	7.6	1020	B86414	hypothetical prote
26	79.5	7.6	1136	T26953	hypothetical prote
27	79.5	7.6	1146	C83304	hypothetical prote
28	79.5	7.6	1557	T23132	hypothetical prote
29	79	7.5	684	T F85075	hypothetical prote

ALIGNMENTS

RESULT 1

T34305
hypothetical protein F54E7.7 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T34305
R:Bentley, D.
submitted to the EMBL Data Library, June 1994
A:Description: The sequence of C. elegans cosmid F54E7.
A:Reference number: Z21502
A:Accession: T34305
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-239 <BEN>
A:Cross-references: EMBL:U00067; PIDN:AAC77519.1; GSPDB:GN00021; CESP:F54E7.7
A:Experimental source: strain Bristol N2; clone F54E7
C:Genetics:
A:Gene: CESP:F54E7.7
A:Map position: 3
A:Introns: 40/3; 78/2; 165/2; 205/1

Query Match 33.9%; Score 355; DB 2; Length 239;
Best Local Similarity 41.0%; Pred. No. 2.3e-24;
Matches 75; Conservative 36; Mismatches 69; Indels 4; Gaps 3;

QY	8	DLPSATIACHLDP	RVFV	DG	LC	RA	KF	ES	FR	YD	KD	T	T	F	Q	P	K	S	K	R	V	R	I	N	S	N	L	S	A	A	67		
DB	27	DLPNAIIVTQ	VP	ED	YD	NK	QD	K	A	N	F	S	S	L	T	Q	E	K	D	I	H	F	D	L	R	S	F	R	R	V	I	86	
QY	68	RRLHKT	EFLG	KML	YFAQ	T	LH	I	-	-	G	S	H	L	A	P	N	P	D	K	O	F	L	I	S	P	P	A	S	P	P	125	
DB	87	KLIVQ	F	S	F	K	G	H	E	L	K	A	F	F	A	Q	I	Y	M	S	A	N	S	O	M	L	S	P	P	L	E	146	
QY	126	VI	NY	D	L	L	A	I	A	I	S	K	L	G	E	K	V	E	L	H	A	T	D	P	P	S	V	V	H	V	C	S	184
DB	147	WCN	F	D	L	M	A	R	L	A	S	F	A	I	D	E	K	I	E	V	H	G	E	L	T	P	A	I	V	H	P	C	205
QY	185	IQT	187																														
DB	206	LQS	208																														

RESULT 2

T49641
nebula related protein [imported] - Neurospora crassa
N:Alternate names: protein B5022.230
C:Species: Neurospora crassa
C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
C:Accession: T49641
R:Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Partmann, B.; Holland, R.; Nyakatura
submitted to the Protein Sequence Database, May 2000
A:Reference number: Z25022

Db 301 EQAPSPAAACEKGEQHSQAEELGPGQEAEDPEEKAVAVPTPPVSPVSTPEVAPPEQ 160
QY 157 -----VVHVCSQDQNEEEEMERMKRPKPKIOTRRPEYTPIH 196
Db 361 LSEAALKAMEEAAVQLEODORHLSKQ-EKMQQLREKLQCEEEIEILRLH 411
RESULT 6
D84485
probable retroelement pol polyprotein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: D84485
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.B.; Umayam, L.; Tallon, L.;
eues, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: D84485
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1274 <STO>
A:Cross-references: GB:AE002093; NID:94309763; PIDN:AAI15532.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g07400
A:Map position: 2

Query Match 8.7%; Score 91.5; DB 2; Length 1274;
Best Local Similarity 23.7%; Pred. No. 7.2;
Matches 51; Conservative 30; Mismatches 83; Indels 51; Gaps 9;
QY 24 VDGLCRAKFESLFTYDKDTTFQYFKSKRVIRNFNSPLSAADAR-LRLHKTFLGKE-- 80
Db 673 IESRVEAKFEGRTGSIENDV-----KQKEQLKAIADKSSSVYRDWYLAKTQPTQODN 727
QY 81 MKLYFAOTLHI-----GSHLAPNPD-----KQFLISPPA-----SPPVGWK 118
Db 728 PKVQTQTPDPVKPTNPATSPPPSKQADVKKETLNELIQPKPGRGRKPSQQPK 787
QY 119 QVEDATPVINDLLYATSKLPGKGYELHA-----ATDTPSVVVHVCSQDQNEEEEEE 173
Db 788 KVSPTMPKXI-----TIKKLQSKQAEKAEEDSDVDVTDKVVSEYNEMLPDEDEEEE 841
QY 174 MERMKRPKPKIOTRR-----PEYTPIHLS 198
Db 842 TERLKSVEIRLKTVKLAPDGRSLINPEAMPHTS 876

RESULT 7
B48516
surfactant protein A receptor - pig
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 07-Apr-1994
C:Accession: B48516
R;Strayer, D.S.; Yang, S.; Jerng, H.H.
J. Biol. Chem. 268, 18679-18684, 1993
A:Title: Surfactant protein A-binding proteins. Characterization and structures.
A:Reference number: A48516; MUID:93366778; PMID:8360162
A:Accession: B48516
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-284 <STR>
A:Cross-references: GB:L10124

Query Match 8.4%; Score 88; DB 2; Length 284;
Best Local Similarity 24.1%; Pred. No. 2.3;
Matches 40; Conservative 24; Mismatches 62; Indels 40; Gaps 7;
QY 27 LCRAKFESLFTYDKDTTFQYFKSKRVIRNFNSPLSAADARLRLHKTFLGKEMKLYEA 86
Db 3 MCYGMGSLFRSGSRQTUFA-----SQMRYAD-----LYAASFIN-----LLYYP 43

QY 87 QTLHGSSHLAPPNP-----DKQFLISPPAS---PPVGWKQVEDATPVINDLLYAI 135
Db 44 FSYLFAARHVLMPHSTVEHTHVDINEMESPLATNRTSVDFKD-----TDYKRHQLTRSI 99
QY 136 SKLGGEXGYELHAATDPTPSVVVVHVCSQDQNEEEEMERMKRPK 181
Db 100 SEIKPPNLPPL-----APQEI THCHDEDDDEEEEEECECGKPK 139

RESULT 8
S29796
hypothetical protein 2280 - evening primrose chloroplast (fragment)
C:Species: chloroplast Oenothera villaricae (evening primrose)
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: S29796; S19986
R;Nimzyk, R.; Schoendorf, T.; Hachtel, W.
Curr. Genet. 23, 265-270, 1993
A:Title: In-frame length mutations associated with short tandem repeats are located in
A:Reference number: S29795; MUID:93169690; PMID:8435856
A:Accession: S29796
A:Molecule type: DNA
A:Residues: 1-630 <NIM>
A:Cross-references: EMBL:X64615; NID:g14331; PID:g14332
C:Genetics:
A:Genome: chloroplast
C:Keywords: chloroplast

Query Match 8.4%; Score 88; DB 2; Length 630;
Best Local Similarity 26.2%; Pred. No. 6.2;
Matches 45; Conservative 23; Mismatches 48; Indels 56; Gaps 10;
QY 16 CHLDP-RVVFV-DGLCRAKFESLFTYDKDTTFQYFKSKRVIRNFNSPLSAADARLRLHK 73
Db 81 CSIDPISIVIKNKLCEAGDSLSYKYW-----FELGTSMKKLTI----- 118
QY 74 TFLGKEMKLYFAQTLHIGS---SHLAPPNPDQFLISPPASPPVGWKQVEDATPVINDY 130
Db 119 -----LLYLLTCSAGSIAQDLLSPGPDEQNLT-----SYGLVEN-----DSD 157
QY 131 LLYAISKLGPGKGYELHAA---TDTPSVVVHVCSQDQNEEEE---EEEME 175
Db 158 LVHGLSDIVHG-LLELEGALVGSSPTREEVEGTEEEVEGTEEEVE 208

RESULT 9
S29795
hypothetical protein 2280 - evening primrose (Oenothera picensis subsp. picensis) chlor
C:Species: chloroplast Oenothera picensis subsp. picensis (evening primrose)
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: S29795; S19983
R;Nimzyk, R.; Schoendorf, T.; Hachtel, W.
Curr. Genet. 23, 265-270, 1993
A:Title: In-frame length mutations associated with short tandem repeats are located in
A:Reference number: S29795; MUID:93169690; PMID:8435856
A:Accession: S29795
A:Molecule type: DNA
A:Residues: 1-721 <NIM>
A:Cross-references: EMBL:X64616; NID:g14334; PID:g14335
C:Genetics:
A:Genome: chloroplast
C:Keywords: chloroplast

Query Match 8.4%; Score 88; DB 2; Length 721;
Best Local Similarity 26.2%; Pred. No. 7.3;
Matches 45; Conservative 23; Mismatches 48; Indels 56; Gaps 10;
QY 16 CHLDP-RVVFV-DGLCRAKFESLFTYDKDTTFQYFKSKRVIRNFNSPLSAADARLRLHK 73
Db 81 CSIDPISIVIKNKLCEAGDSLSYKYW-----FELGTSMKKLTI----- 118
QY 74 TFLGKEMKLYFAQTLHIGS---SHLAPPNPDQFLISPPASPPVGWKQVEDATPVINDY 130
Db 119 -----LLYLLTCSAGSIAQDLLSPGPDEQNLT-----SYGLVEN-----DSD 157

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 14, 2002, 18:14:56 ; Search time 1655.79 Seconds
(without alignments)
3480.130 Million cell updates/sec

Title: US-09-782-953-3

Perfect score: 1048

Sequence: 1 MEEVDLQDLPSATIACHLDP.....RPKXLIQTRRPEYTPHLS 198

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DBV=xlh
-Q/cgn2_1/USPTO.spool/US09782953/runat 11122002 114429 17340/app query.fasta_1.1173
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPECL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=oct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODES=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09782953 -CGCN 1 2844 @runat 11122002 114429 17340 -NCPU=6 -ICPU=3
-NO XUPXY -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl :

1: gb_ba :
2: gb_htg :
3: gb_in :
4: gb_om :
5: gb_ov :
6: gb_pat :
7: gb_ph :
8: gb_pl :
9: gb_pr :
10: gb_ro :
11: gb_ste :
12: gb_sy :
13: gb_un :
14: gb_vi :
15: em_ba :
16: em_fun :
17: em_hum :
18: em_in :
19: em_mu :
20: em_om :
21: em_ox :
22: em_ov :
23: em_pat :
24: em_ph :
25: em_pi :
26: em_ro :
27: em_sts :
28: em_un :

29: em_vi :
30: em_htg_hum :
31: em_htg_inv :
32: em_htg_other :
33: em_htg_mus :
34: em_htg_pln :
35: em_htg_rod :
36: em_htg_mam :
37: em_htg_vrt :
38: em_sy :
39: em_hcgo_hum :
40: em_hcgo_mus :
41: em_hcgo_other :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1048	100.0	597	6	AX365312 Sequence
2	1048	100.0	597	10	AF237789 Mus muscu
3	1040	99.2	621	10	AF282255 Mus muscu
4	1040	99.2	2125	10	AF260717 Mus muscu
5	1002.5	95.7	2212	6	AX365324 Sequence
6	1002.5	95.7	2227	9	HSU85266
7	1002.5	95.7	2289	9	BC002864 Homo sapi
8	1002.5	95.7	2407	9	AK092184 Homo sapi
9	920.5	87.8	597	6	AX365315 Sequence
10	920.5	87.8	597	10	AF237790 Mus muscu
11	911	86.9	2141	10	AF263239 Mus muscu
12	911	86.9	2198	10	BC033551 Mus muscu
13	890	84.9	2224	10	AF263240 Mus muscu
14	887.5	84.7	626	10	AB075973 Rattus no
15	884.5	84.4	2216	10	CGU60263 Cricetulus
16	872.5	83.3	2331	6	AX365321 Sequence
17	872.5	83.3	2346	9	HSU85267 Homo sapien
18	859.5	82.0	2348	6	AX281651 Sequence
19	847.5	80.9	2173	6	AX410694 Sequence
20	847.5	80.9	2173	9	HSU28833 Homo sapien
21	847.5	80.9	2284	9	HSU85265 Homo sapien
22	847.5	80.6	2679	9	AK055845 Homo sapi
23	844.5	80.6	2174	6	AR034241 Sequence
24	839.5	80.1	798	9	AF400429 Homo sapi
25	749.5	71.5	599	6	AX365311 Sequence
26	709	67.7	562	9	HSU53821 Homo sapien
27	627.5	59.9	3295	10	AB061525 Mus muscu
28	620.5	59.2	934	6	AX074351 Sequence
29	620.5	59.2	3159	6	AX420425 Sequence
30	620.5	59.2	3261	9	AK090990 Homo sapi
31	613.5	58.5	594	6	AX365318 Sequence
32	613.5	58.5	594	10	AF237791 Mus muscu
33	613.5	58.5	659	10	AF237887 Mus muscu
34	608.5	58.1	3227	10	AB061524 Mus muscu
35	607.5	58.0	3240	9	AY034085 Homo sapi
36	607.5	58.0	3253	9	AY034086 Homo sapi
37	597	57.0	776	10	AF237888 Mus muscu
38	596.5	56.9	3184	6	AX329596 Sequence
39	596.5	56.9	3184	6	AX333015 Sequence
40	596.5	56.9	3184	6	AX365327 Sequence
41	596.5	56.9	3184	6	AX420436 Sequence
42	596.5	56.9	3184	9	HUMZAK14
43	596	56.9	828	6	AX365330
44	596	56.9	828	6	AX420437 Sequence
45	596	56.9	828	9	AF176116 Homo sapi

ALIGNMENTS

RESULT 1

AX365312	AX365312	597 bp	DNA	linear	PAT 15-FEB-2002
LOCUS	Sequence 2 from Patent WO0204491.				
DEFINITION	AX365312				
ACCESSION	AX365312				
VERSION	AX365312.1	GI:18697043			
KEYWORDS					
SOURCE	house mouse.				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
AUTHORS	Williams S.R. and Rothgamel B.				
TITLE	Methods and compositions relating to muscle selective calcineurin interacting protein (mciip)				
JOURNAL	Patent: WO 0204491-A 2 17-JAN-2002; Patent of Regents, The University of Texas System (US) ; Williams, Sanders R. (US) ; Rothgamel, Beverly (US)				
FEATURES	Location/Qualifiers				
Source	1..597				
	/organism="Mus musculus"				
	/db_xref="taxon:10090"				
	1..597				
CDS	/note="unnamed protein product"				
	/codon_start=1				
	/protein_id="CAD23808.1"				
	/db_xref="GI:18697044"				
	/translation="MEVDLQDIPSATIACHDHPVPVGLCRAKFSLPFTYDKDTT FOYKSFKRVKRNFSNPASADARLRKTELEGKMLTYPAQTIHSSHLAPNP KQPLSPSPSPVGVKQVEDATPVINQDLVAISLGGKRELMADTPSPVYVH CESDQNEEBEERMRKPKPKI1QTRRPEYPIHLS"				
BASE COUNT	160 a 170 c 151 g 116 t				
ORIGIN					
Alignment Scores:					
Pred. No.:	5, 06e-96	Length:	597		
Score:	1048.00	Matches:	198		
Percent Similarity:	100.00%	Conservative:	0		
Best Local Similarity:	100.00%	Mismatches:	0		
Query Match:	100.00%	Indels:	0		
DB:	6	Gaps:	0		
US-09-782-953-3 (1-198) x AX365312 (1-597)					
QY	1 MetGUGUValAaPLeuGInAaPLeuPProSerAlaThrIleAlaCyHisLeuAaPPro	20			
Db	1 ATGGAGGAGGTGGATTGCGAGGACCTGGCGAGCGCCACCATGCGCCACCTGGACCGG	60			
QY	21 ArgValAPheValAaPGLyLeuCyArgAlaLysPheGluSerLeuPheArgThrTyArgP	40			
Db	61 CGCGTTCGTGGAGCGCGCTGTGCGGGCCAAATTTGAATCCCTCTTCAGAACATATGAC	120			
QY	41 LysAaPThrThrPheGlnTyPheLysSerPheLysArgValArgIleAaPheSerAaP	60			
Db	121 AAGGACACCACTTCCAGTATTTTAAAGAGCTTAAACGTCCTCGGATTAACCTTCAGCAAC	180			
QY	61 ProLeuSerAlaAlaAaPAlaArgLeuArgLeuHisLysThrGluPheGluGlyGlu	80			
Db	181 CCTTATTCGACACCGCATGCCAGGCTGCGGCTGCACAAACGAGTTCTCGGGAGAGAA	240			
QY	81 MetLysLeuTyPheAlaGlnThrLeuHisIleGlySerSerThrIleuAlaProProAaP	100			
Db	241 ATGMAATGTGATTTTGTCTCAGACTTTTACATATGAGAAAGTTACACCTGGCTCGCCCAAT	300			
QY	101 ProAaPLeuGlnPheLeuLysSerProAlaLysSerProArgValGlyTPLYGlnVal	120			
Db	301 CCCGACAAACAGTTCCTCATCTCCCTCCGCGCTCTCTCCCTGGCTGGAAACAAGTA	360			
QY	121 GluAaPAlaThrProValIleAaPThrAaPLeuLeuTyAlaIleSerLysLeuGlyPro	140			
Db	361 GAGAGTCCGACCCCGCATTAATTAAGATCTTTATATGCCATCTCCAAAGCTGGGGCCA	420			
QY	141 GlyGluLysTyGluLeuHisAlaAlaThrAaPProThrProSerValValValHisVal	160			

Db	421	GGAGGAGAAGTATGAACACTGCATGCAGCGACAAGACCCTCCCACTGTGGTGTCACAGTG	480
Oy	161	CysGIuserAapGIngluaenGlucIuglucIugluCmetGLuaGmelVayarPro	180
Db	481	TGTGAGAGTGCACCAAGAAATGAGAGAGAAAGAGAAAGATGTGAGAGAAATGAAAGAGCCC	540
Oy	181	LysPpolySelleleGINthRaxrArxProGluYurThrPhroLeHslauSer	198
Db	541	AAGCCCAAATCATTCACAGACGAGAGACCGAGTAGTACACCGATCCACTTAGC	594
<hr/>			
RESULT 2			
LOCUS	AF237789	597 bp	RNA linear ROD 12-APR-2000
DEFINITION	Mus musculus myocyte-enriched calcineurin interactin protein 1		
ACCESSION	AF237789		
VERSION	AF237789.1		
KEYWORDS	splice variant 1 mRNA, complete cds.		
SOURCE	Mus musculus.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathii; Muridae; Murinae; Mus		
TITLE	Rocheamel,B.A., Vega,R.B., Yang,J., Wu,H., Baesel-Duby,R.S. and Williams,R.S.		
JOURNAL	A Protein Encoded within the Down Syndrome Critical Region is Enriched in Striated Muscles and Inhibits Calcineurin Signaling		
REFERENCE	J. Biol. Chem. (2000) In press		
AUTHORS	2. (bases 1 to 597)		
TITLE	Rocheamel,B.A., Vega,R.B., Yang,J., Wu,H., Baesel-Duby,R.S. and Williams,R.S.		
JOURNAL	Direct Submission		
REFERENCE	Submitted (22-FEB-2000) Internal Medicine, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, TX 75390, USA		
FEATURES	Location/Qualifiers		
Source	1..597 /organism="Mus musculus"		
	/dd_xref="taxon:10090"		
	/chromosome="16"		
CDS	1..597 /note="similar to Homo sapiens DSCR1, MCIPI, similar to EST sequence of Genbank Accession Number AA734360"		
	/codon_start=1		
	/product="myocyte-enriched calcineurin interactin protein 1 splice variant 1"		
	/protein_id="AAF63485.1"		
	/db_xref="GI:7542526"		
	/translation="MEEVDLDDPSATIACHLDPRVEVDGLCRKAFESLPFTYDQTTITFOYKSPFRVRINFSNPLSADARLRHKTEFLGEKKMLYAQRLHGSSHLAPNDPKOFLSPASPPVKMKVEDPTPIYNVLDAISKIDRGEXELHAANDTPRSVVVHHCESDNEEEEEEREMRKRPKIIIQRRREPYTHLS"		
BASE COUNT	160 a 170 c 151 g 116 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	5,06e-96	Length:	597
Score:	1048.00	Matches:	198
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatch:	0
Query Match:	100.00%	Indels:	0
DB:	10 Gaps:	0	
<hr/>			
US-09-782-953-3 (1-198) x AF237789 (1-597)			
Oy	1	MetGLuGlUvaLaSplEuGlInaSplEuProSeRaLaThrILaALCyehIsLeuAspPro	20
Db	1	ATGGAGGAGGTGATTGTCAGAGACTGCCGAGACGCCACATGCGCTCGCACCTGGACCCG	60
Oy	21	ArgValaPheValaSpGlyLeuCYaRgaLaLyaphneGluSerLeuPheArgThrTyaaP	40
Db	61	CGCGTGTTCGTGGACGGCTGTGCCGGGCCCAATTGTAATCCCTCTTCAGAACATATGAC	120

QY 41. LysAspThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAen 60
 Db 121 AAGGACACCACTTCAGTATTTAAGAGCTTCAACAGTGTCCGGATAAATTCAGCAAC 180
 QY 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
 Db 181 CCTTATCTGCAGCCGATGCCAGCTGCCGCTGCACAAAGCCAGTTCCTGGGAAGGAA 240
 QY 81 MetLysLeuTyrPheAlaGlnThrLeuHisLysGlySerSerHisLeuAlaProAen 100
 Db 241 ATCAAGTGTGATTTGCTCAGACTTTACACATAGGAAGTTCACACCTGCTCGGCCAAT 300
 QY 101 ProAspLysGlnPheLeuLeuSerProProAlaSerProProValGlyTyrLysGlnVal 120
 Db 301 CCCGACAAACAGTTCCTCATCTCCCTCCGCGCTCTCTCCCTCCGCTGGAAACAAGTA 360
 QY 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro 140
 Db 361 GAAGATGCCACCCCGCTCAATAATACGATCTTTATATGCCATCTCCAGCTGGGGCCA 420
 QY 141 GlyGlyLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160
 Db 421 GGAGAGAAGTATGAATGCATGCAGCAGCAGACCCCACTCCCACTGTGTGTGTCCAGTG 480
 QY 161 CysGluSerAspGlnGluAsnGluGluGluGluGluGluMetGluArgMetLysArgPro 180
 Db 481 TGTGAGAGTGACCAAGAGAATGAGGAGGAGAGAGAGATGAGAGAATGAGAGACCC 540
 QY 181 LysProLysIleLeuGlnThrArgArgProGluTyrThrProIleHisLeuSer 198
 Db 541 AAGCCCAAAATCATCCAGACACGAGACCGAGTACACCCGATCCACCTTAGC 594

RESULT 3
 AF282255 621 bp mRNA linear ROD 17-NOV-2000
 LOCUS AF282255 Mus musculus Down syndrome candidate region 1 protein (Dscr1) mRNA, complete cds.

ACCESSION AF282255
 VERSION AF282255.1 GI:9652249

KEYWORDS Mus musculus.
 SOURCE Mus musculus.

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 621)
 Strippoli, P., Petrini, M., Lenzi, L., Carinci, P. and Zannotti, M.

TITLE The murine DSCR1-like (Down syndrome candidate region 1) gene family: conserved synteny with the human orthologous genes

JOURNAL Gene 257 (2), 223-232 (2000)
 MEDLINE 20534792
 PUBMED 11080588

REFERENCE 2 (bases 1 to 621)
 Strippoli, P., Petrini, M., Lenzi, L., Carinci, P. and Zannotti, M.

AUTHORS Direct Submission
 TITLE Submitted (24-JUN-2000) Istituto di Istologia ed Embriologia Generale, Universite di Bologna, Via Belmeloro, 8, Bologna, BO 40136, Italy

FEATURES Location/Qualifiers
 1..621
 /organism="Mus musculus"
 /strain="BALB/c"
 /db_xref="taxon:10090"
 /chromosome="16"
 /sex="male"
 /tissue_type="brain"
 1..621
 /genes="Dscr1"
 8..604
 /genes="Dscr1"
 /note="similar to the Homo sapiens Down syndrome candidate region 1 protein"
 /codon_start=1
 /product="Down syndrome candidate region 1 protein"

source
 1..621
 /organism="Mus musculus"
 /strain="BALB/c"
 /db_xref="taxon:10090"
 /chromosome="16"
 /sex="male"
 /tissue_type="brain"
 1..621
 /genes="Dscr1"
 8..604
 /genes="Dscr1"
 /note="similar to the Homo sapiens Down syndrome candidate region 1 protein"
 /codon_start=1
 /product="Down syndrome candidate region 1 protein"

Gene
 CDS

/protein_id="AAF91461.1"
 /db_xref="GI:9652250"
 /translation="MEEVDLQDLPSATIACHLDPRVFDGLCKAFPSLFRFYDKDTT
 KQFLSPKRVIRPNFNSPLSADARLRHKTFLGKEMKLYPAOTLHIGSSSLAPNP
 KQFLSPSPPPVGMKQVEDATPVINDLLVLAISKLGPEKVELHAATDTTPSVVVHV
 CESDOENESEEMERMKRPKPKIIQTRPEYTPILHS"
 BASE COUNT 166 a 179 c 159 g 117 t
 ORIGIN

Alignment Scores:
 Pred. No.: 3,35e-95 Length: 621
 Score: 1040.00 Matches: 197
 Percent Similarity: 99.49% Conservative: 0
 Best Local Similarity: 99.49% Mismatches: 1
 Query Match: 99.24% Indels: 0
 DB: 10 Gaps: 0

US-09-782-953-3 (1-198) x AF282255 (1-621)

QY 1 MetGluGluValAspLeuGlnAspLeuProSerAlaThrIleAlaCysHisLeuAspPro 20
 Db 8 ATGGAGGAGGTGGATCTGCAGGACCTGCAGCGCCAGCATCGCTGCACCTGACCCG 67

QY 21 ArgValPheValAspGlyLeuCysArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40
 Db 68 CGCGTGTTCGTGGAGCGGCTGTGCCGGGCCAAATTTGAATCCCTCTTCAGAAACATATGAC 127

QY 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAen 60
 Db 128 AAGGACACCACTTCAGTATTTAAGAGCTTCAACAGTGTCCGGATAAATTCAGCAAC 187

QY 61 ProLeuSerAlaAlaAspAlaArgLeuHisLysThrGluPheLeuGlyLysGlu 80
 Db 188 CCTTATCTGCAGCGATGCCAGCTGCCGCTGCACAGACCGAGTTCCTGGGAAGGAA 247

QY 81 MetLysLeuTyrPheAlaGlnThrLeuHisLysGlySerSerHisLeuAlaProAen 100
 Db 248 ATGAAGTGTGATTTTGTCTCAGACTTTACACATAGGAAGTTCACACCTGCTCGGCCAAT 307

QY 101 ProAspLysGlnPheLeuLeuSerProProAlaSerProProValGlyTyrLysGlnVal 120
 Db 308 CCCGACAAACAGTTCCTCATCTCCCTCCGCGCTCTCTCCCTGGCTGGGAACAGTA 367

QY 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro 140
 Db 368 GAAGATGCCACCCCGTCATAAATTCAGATCTTTTATATGCCATCTCCAAGCTGGGCCA 427

QY 141 GlyGlyLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160
 Db 428 GGAGAGAAGTATGAACCTGCATGCAGCAGACACCACTCCCACTGTGTGTGTCCACGTTG 487

QY 161 CysGluSerAspGlnGluAsnGluGluGluGluGluMetGluArgMetLysArgPro 180
 Db 488 TGTGAGAGTGACCAAGAGATGAGGAGGAGAGAGATGAGAGATGAGAGATGAGAGACCC 547

QY 181 LysProLysIleLeuGlnThrArgArgProGluTyrThrProIleHisLeuSer 198
 Db 548 AAGCCCAAAATCATCCAGACACGAGACCGGAGTACACACCACTCCACCTCAGC 601

RESULT 4
 AF260717 2125 bp mRNA linear ROD 22-MAR-2001
 LOCUS AF260717 Mus musculus Down syndrome candidate region 1 (Dscr1) mRNA, complete cds.

DEFINITION AF260717
 ACCESSION AF260717
 VERSION AF260717.1 GI:7839596

KEYWORDS Mus musculus.
 SOURCE Mus musculus.

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 2125)
 Casas, C., Martinez, S., Pritchard, M.A., Fuentes, J.J., Nadal, M.,

Guinera, J., Arbones, M., Florez, J., Soriano, E., Estivill, X. and Alcantara, S.
 Descl, a novel endogenous inhibitor of calcineurin signaling, is expressed in the primitive ventricle of the heart and during neurogenesis.
 Mech. Dev. 101 (1-2), 289-292 (2001)

JOURNAL MEDLINE
 21152920
 PUBLISHED 11231093
 2 (bases 1 to 2125)
 REFERENCES
 Fuentes, J., Pritchard, M., Pucharcas, C. and Estivill, X.
 Down syndrome candidate region 1 (Descl), one of three alternatively spliced exon 1
 Unpublished
 3 (bases 1 to 2125)
 REFERENCES
 Fuentes, J., Pritchard, M., Pucharcas, C. and Estivill, X.
 Direct Submission
 Submitted (26-APR-2000) Centre de Genetica Medica i Molecular, IRO, Avia, Castelldefels Km. 2,7, L'Hospital de Llobregat, Barcelona 08907, Spain

FEATURES
 source
 1. .2125
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /chromosome="16"
 1. .2125
 /gene="Descl"
 15. .611
 /gene="Descl"
 /note="calcineurin inhibitor"
 /codon_start=1
 /product="Down syndrome candidate region 1"
 /protein_id="AA07043.1"
 /db_xref="GI:7839597"
 /translation="MEVDLDLPSATIACHLDPFVVDGLCRKFESLFRPYDKDTT
 FOYKSKFRVIRINFSNPLSADARLKHTEFLGEMKLYFAOTLHGSSHLAPNPD
 KQFLSPSPAPVGMKQVEDATPVINDLYLSLGGGEKELHAADTTTPSVVHV
 CESDQNEEEREMERMRPKKIIQTRRPETPIHLS"

BASE COUNT 487 a 536 c 563 g 536 t 3 others
 ORIGIN

Alignment Scores:
 Pred. No.: 1,47e-94 Length: 2125
 Score: 1040.00 Matches: 197
 Percent Similarity: 99.49% Conservative: 0
 Best Local Similarity: 99.49% Mismatches: 1
 Query Match: 99.24% Indels: 0
 DB: 10 Gaps: 0

US-09-782-953-3 (1-198) x AF260717 (1-2125)

OY 1 MetGluGluValAspLeuGlnAspLeuProSerAlaThrIleAlaCyHHisLeuAspPro 20
 |||||
 DB 15 ATGAGAGAGATGATCTGACAGACCTGCCAGGCGCACCATCGCTGCACCTGAGCCCG 74
 |||||
 OY 21 ArgValAlpheValAspGlyLeuCyArgAlaLysPheGluSerLeuPheArgThrTyraap 40
 |||||
 DB 75 CCGGTGTGTGTGACCGCTGTGCGCGCCAAATTTGAATCCCTCTTGAGAACATATGAC 134
 |||||
 OY 41 LysAspThrThrPheGlnTyPheLysSerPheLysArgValArgIleAsnPheserAsn 60
 |||||
 DB 135 AAGGACACCACTTCCAGATTTTAAAGACTTCAACGTCGTCCGATTAACCTTCAGCAAC 194
 |||||
 OY 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
 |||||
 DB 195 CCTTATCTGACGACCGCATCCAGGCTGCGGCTGCACAAAGACCGAGTTCCTGGGAAAGAA 254
 |||||
 OY 81 MetLysLeuTyPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProProAsn 100
 |||||
 DB 255 ATGAAGTTGATTTTGTCTAGACTTACATAGGAAGTTTCAACCTGCTCGGCCCAAT 314
 |||||
 OY 101 ProAspLysGlnPheLeuLysSerProAlaSerProProValGlyTrpLysGlnVal 120
 |||||
 DB 315 CCGGCAAAACAATTCTCTATCTCCCTCCGCGCTCTCTCCGTTGGCTGGAAACAAGTA 374
 |||||

OY 121 GluAspAlaThrProValIleAsnTyPheAspLeuTyPheAlaIleSerLysLeuGlyPro 140
 |||||
 DB 375 GAAGATGCCACCCCGCTGATTAATACGATCTTTATATGTCATCTCCAGCTGGGCCCA 434
 |||||
 OY 141 GlyGluValTyPheGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160
 |||||
 DB 435 GGAGAGAGTATGAATGATGATGACAGACGACACCTCCAGTGTGTGTCCACGCG 494
 |||||
 OY 161 CyGluSerAspGlnGluAsnGluGluGluGluGluGluGluGluGluGluGluGluGlu 180
 |||||
 DB 495 TTGTAGAGTACCAAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGATGAGAGAGACC 554
 |||||
 OY 181 LysProLysIleIleGlnThrArgArgProGluTyPheThrProIleHisLeuSer 198
 |||||
 DB 555 AAGCCCAAAATATCATCCAGACGAGAGACCGGAGTACACCATCCTCAGCC 608
 |||||

RESULT 5
 AX365324
 LOCUS AX365324 2212 bp DNA linear PAT 15-FEB-2002
 DEFINITION Sequence 14 from Patent WO0204491.
 ACCESSION AX365324
 VERSION AX365324.1 GI:18697051
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 1 Williams, S.R. and Rothermel, B.
 Methods and compositions relating to muscle selective calcineurin
 interacting protein (mcip)
 Patent: WO 0204491-A 14 17-JAN-2002;
 Board of Regents, The University of Texas System (US); Williams,
 Sanders R. (US); Rothermel, Beverly (US)

FEATURES
 source
 1. .2212
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 25. .618
 /note="named protein product"
 /codon_start=1
 /protein_id="CAD23812.1"
 /db_xref="GI:18697052"
 /translation="MEVDLDLPSATIACHLDPFVVDGLCRKFESLFRPYDKDTT
 FOYKSKFRVIRINFSNPLSADARLKHTEFLGEMKLYFAOTLHGSSHLAPNPD
 KQFLSPSPAPVGMKQVEDATPVINDLYLSLGGGEKELHAADTTTPSVVHV
 CESDQNEEEREMERMRPKKIIQTRRPETPIHLS"

BASE COUNT 588 a 455 c 529 g 640 t
 ORIGIN

Alignment Scores:
 Pred. No.: 8,74e-91 Length: 2212
 Score: 1002.50 Matches: 191
 Percent Similarity: 97.98% Conservative: 3
 Best Local Similarity: 96.46% Mismatches: 3
 Query Match: 95.66% Indels: 1
 DB: 6 Gaps: 1

US-09-782-953-3 (1-198) x AX365324 (1-2212)

OY 1 MetGluGluValAspLeuGlnAspLeuProSerAlaThrIleAlaCyHHisLeuAspPro 20
 |||||
 DB 25 ATGAGAGAGTGAACCTTCAGACCTGCCAGCGCACCATCGCTGCACCTGAGCCCG 84
 |||||
 OY 21 ArgValAlpheValAspGlyLeuCyArgAlaLysPheGluSerLeuPheArgThrTyraap 40
 |||||
 DB 85 CCGGTGTGTGTGACCGCTGTGCGCGCCAAATTTGATGCTCTTAAAGACGATATGAC 144
 |||||
 OY 41 LysAspThrThrPheGlnTyPheLysSerPheLysArgValArgIleAsnPheserAsn 60
 |||||
 DB 145 AAGGACATCACTTTCAGTATTTTAAAGCTTCAACGAGTCAAGATTAACCTTCAGCAAC 204
 |||||

QY 61. ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
Db 205 CCCTCTCCGAGAGATGCCAGGCTCCAGCTGATGAAGACTGAGTTCTTGGGAAGGAA 264
QY 81 MetLysLeuTyrPheAlaGlnThrLeuHisLysGlySerSerHisLeuAlaProProAsn 100
Db 265 ATGAAGTTATATTTTGTCTAGACCTTACACATAGGAAGCTCACACCTGGCTCCGCCAAAT 324
QY 101 ProAspLysGlnPheLeuLysSerProProAlaSerProProValGlyTyrLysGlnVal 120
Db 325 CCAGACAAGCAGTTCTTGATCTCCCTCCGCTCTCCGCTCCGCTCCGCTCCGCTCCGCT 384
QY 121 GluAspAlaThrProValLysLeuAsnTyrAspLeuLeuTyrAlaLysSerLysLeuGlyPro 140
Db 385 GAAGATCGACCCCGAGCTATGAATCTATGATCTCTATATGCTATCTCAAGCTGGGGCCA 444
QY 141 GlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160
Db 445 GGGGAAAAGTATGAATTTGCACGAGCAGCTGACACCACTCCAGCTGGTGGTCCATGTA 504
QY 161 CysGluSerAspGlnGluAsnGluGluGluGluMetGluArgMetLysArgPro 180
Db 505 TGTGAGAGTGATCAAGAG---AAGGAGGAGAGAGGAAATGGAAAGATGAGAGACCT 561
QY 181 LysProLysLysLysGlnThrArgArgProGluTyrThrProLysHisLeuSer 198
Db 562 AAGCCAAAATTTATCCAGACCGAGGAGCGGAGTACACGCGGATCCACCTCAGC 615

RESULT 6
HSU85266 2227 bp mRNA linear PRI 19-APR-2000
LOCUS Homo sapiens down syndrome candidate region 1 (DSCR1) gene,
DEFINITION alternative exon 1, complete cds.
ACCESSION U85266
VERSION U85266.2 GI:7596913
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2227)
AUTHORS Fuentes,J.J., Pritchard,M.A., Planas,A.M., Bosch,A., Ferrer,I. and Estivill,X.
TITLE A new human gene from the Down syndrome critical region encodes a proline-rich protein highly expressed in fetal brain and heart
JOURNAL Hum. Mol. Genet. 4 (10), 1935-1944 (1995)
MEDLINE 96121593
PUBMED 8595418
REFERENCE 2 (bases 1 to 2227)
AUTHORS Fuentes,J.J., Pritchard,M.A. and Estivill,X.
TITLE Genomic organization, alternative splicing, and expression patterns of the DSCR1 (Down syndrome candidate region 1) gene
JOURNAL Genomics 44 (3), 358-361 (1997)
MEDLINE 97468152
PUBMED 9325060
REFERENCE 3 (bases 1 to 2227)
AUTHORS Fuentes,J.J., Pritchard,M. and Estivill,X.
TITLE Direct Submission
JOURNAL Submitted (14-JAN-1997) Genetica Molecular, Institut de Recerca Oncologica, Autovia de Castelldefels, km 2.7, Hospitalet, Barcelona 08907, Spain
REFERENCE 4 (bases 1 to 2227)
AUTHORS Fuentes,J.J., Pritchard,M. and Estivill,X.
TITLE Direct Submission
JOURNAL Submitted (19-APR-2000) Genetica Molecular, Institut de Recerca Oncologica, Autovia de Castelldefels, km 2.7, Hospitalet, Barcelona 08907, Spain
REMARK Sequence update by submitter
COMMENT On Apr 19, 2000 this sequence version replaced gi:2612865.
Fuentes, J.J. Hum.Mol.Genet. 4, 1935-1944, 1995.
FEATURES
Location/Qualifiers
1..2227
/organism="Homo sapiens"

/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.1-q22.2"
1..2227
/gene="DSCR1"
25..618
/gene="DSCR1"
/note="down syndrome candidate region 1; one of four alternatively spliced exon 1"
/codon_start=1
/protein_id="AAB84371.2"
/db_xref="GI:7596914"
/translation="MEVDLQDLPSATIACHLDPRFVFDGLCKRAKPFSLRTYDKDITFOYFKSFRVRINFSNPFSAADARIQLKHELMKLYFACTLHGSSHLAPNPDKQFLISPPASPVGKHQVEDATPVINYDLLVAISKLGPEKELYELHATDTTTSVVVHV CESDQKEEEEEEMERMRPKLIITRRPEYTPPIHLS"
BASE COUNT 603 a 455 c 529 g 640 t
ORIGIN
Alignment Scores:
Pred. No.: 8 81e-91 Length: 2227
Score: 1002.50 Matches: 191
Percent Similarity: 97.98% Conservative: 3
Best Local Similarity: 96.46% Mismatches: 3
Query Match: 95.66% Indels: 1
DB: 9 Gaps: 1
US-09-782-953-3 (1-198) x HSU85266 (1-2227)
QY 1 MetGluGluValAspLeuGlnAspLeuProSerAlaThrIleAlaCysHisLeuAspPro 20
Db 25 ATGGAGGAGTGGACCTGCAGGACCTGCCAGGCCACCATCGCTGTACCTGGAGCCG 84
QY 21 ArgValPheValAspGlyLeuCysArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40
Db 85 CGCGTGTTCGTGGACGCGCTGTGCCGGGCCAAATTTAGTCCCTCTTTAGGAGTATGAC 144
QY 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60
Db 145 AAGGACATCACCTTTCAGTATTTTAAAGAGCTTCAACGAGTCAAGATAAATTCACCAAC 204
QY 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
Db 205 CCCTTCTCCGAGCAGATCCAGGCTCCAGCTGCATAGACTGAGTTCTTGGGAAGGAA 264
QY 81 MetLysLeuTyrPheAlaGlnThrLeuHisLysGlySerSerHisLeuAlaProProAsn 100
Db 265 ATGAAGTTATATTTTGTCTAGACCTTACACATAGGAAGCTCACACCTGGCTCCGCCAAAT 324
QY 101 ProAspLysGlnPheLeuLysSerProProAlaSerProProValGlyTyrLysGlnVal 120
Db 325 CCAGACAAGCAGTTCTTGATCTCCCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCT 384
QY 121 GluAspAlaThrProValLysLeuAsnTyrAspLeuLeuTyrAlaLysSerLysLeuGlyPro 140
Db 385 GAAGATCGACCCCGAGCTATGAATTTGCACGAGCAGCTGACACCACTCCAGCTGGTGGTCCATGTA 504
QY 141 GlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160
Db 445 GGGGAAAAGTATGAATTTGCACGAGCAGCTGACACCACTCCAGCTGGTGGTCCATGTA 504
QY 161 CysGluSerAspGlnGluAsnGluGluGluGluMetGluArgMetLysArgPro 180
Db 505 TGTGAGAGTGATCAAGAG---AAGGAGGAGAGAGGAAATGGAAAGATGAGGAGACCT 561
QY 181 LysProLysLysLysGlnThrArgArgProGluTyrThrProLysHisLeuSer 198
Db 562 AAGCCAAAATTTATCCAGACCGAGGAGCGGAGTACACGCGGATCCACCTCAGC 615
RESULT 7
BC002864 BC002864 2289 bp mRNA linear PRI 12-JUL-2001
LOCUS Homo sapiens, Down syndrome critical region gene 1, clone MGC:10263
DEFINITION

IMAGE:3944959, mRNA, complete cds.

BC002864
BC002864.1 GI:12804022

KEYWORDS
MGC.

SOURCE
Homo sapiens.

ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 2289)

REFERENCE
1 Strausberg, R.
Direct Submission
Submitted (05-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nigrl.nih.gov

Shevchenko, Y., Wechterby, K.D., Beckstrom-Sternberg, S.M.,
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legeski, R.,
Lim, M., Maduro, Q.L., Masiello, C., Mastrian, S.D., McCloskey, J.C.,
McDowell, U., Pearson, R., Snyder, B., Stancin, S., Thomas, P.J.,
Tjongson, E.E., Touchman, J.W., Tsurgon, C., Vogt, J.L., Walker, M.A.,
Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
Series: IRAL Plate: 15 Row: a Column: 23
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 7596913.

FEATURES
source
1..2289
Location/Qualifiers

ORIGIN
1..2289
/organism="Homo sapiens"
/db_xref="LOCUSID:1827"
/db_xref="taxon:9606"
/clone="MGC:10263 IMAGE:3944959"
/cfeature="type="Lung, small cell carcinoma"
/clone_id="NIG MGC 7"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
66..659
/codon_start=1
/product="Down syndrome critical region gene 1"
/protein_id="AAH02864.1"
/db_xref="GI:12804023"
/translation="MEVDLQDLPATTAACHLDPVFPVDFGCRAPFESLFRYDIT
FOYRFSKRVIRINFSNPASADARLQHLKTEFLGKEMLYEFQTHLHSSHPANPD
KQFLISPAVPGWQVQEDATPVNIYDLVAISKLGKEXEYLAADTTTPVVVAVH
CESDKEEEMERMRPKPIOTRPPETPHLS"

BASE COUNT
628 a 465 c 555 g 641 t
ORIGIN

Alignment Scores:
Pred. No.: 9.1e-91 Length: 2289
Score: 1002.50 Matches: 191
Percent Similarity: 97.98% Conservative: 3
Best Local Similarity: 96.46% Mismatches: 3
Query Match: 95.66% Indels: 1
DB: 9 Gaps: 1

US-09-782-953-3 (1-198) X BC002864 (1-2289)

1 MetGluGluValAspLeuGlnAspLeuProSerAlaThrTlleAlaCyshleuAspPro 20

Db 66 ATGAGAGAGCTGACCTGACGAGACCTGCCAGCGCCACACACTCGCTGTGACCTGGACCGC 125
Qy 21 ArgValPheValAspGlyLeuCyArgAlaIlePheGluSerLeuPheArgThrTyArgp 40
Db 126 CGCGTGTTCGGACGCGCGCTGTGCGCGGCCCAATTATTAAGTCCCTCTTATGAGACGTATGAC 185
Qy 41 LysAspThrThrPheGlnTyrPheLeuSerPheLeuValArgIleLeuPheSerAsn 60
Db 186 AAGGACATCACCTTTCAGTATTTTAAGAGCTTCAAGAGCTCAAGATTAACCTTACGACAC 245
Qy 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisIleSerThrGluPheLeuGlyLeu 80
Db 246 CCTTCTCCGACGAGAGATGCGAGCTCCAGCTCCAGCTGATAGACTGATTTCTGGAAAGAA 305
Qy 81 MetLeuLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProProAsn 100
Db 306 ATGAAGTTATATTTTGGCTCAGACCTTACACATAGGAAGCTCACACCTGGCTCCCAAT 365
Qy 101 ProAspLeuGlnPheLeuIleSerProProIleSerProProValGlyTyrPheGlnVal 120
Db 366 CCAGCAAGCAGTTTCTGATCTCCCTCCCTCCCTCCGCGAGTGGAGAAACAAAGTG 425
Qy 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerTyLeuGlyPro 140
Db 426 GAAGATGCGACCCAGCTCATATACTATGATCTTATATGCACTCCAGCTGGCGCA 485
Qy 141 GlyGluTyrGluLeuHisAlaIleThrAspProThrProSerValValIleVal 160
Db 486 GGGGAAAGTATGATATTCACGACGCGCTGACACCACTCCAGCTGGTGTGCATCTA 545
Qy 161 CysGluSerAspGlnGluAsnGluGluGluGluGluMetGluAspMetLeuArgPro 180
Db 546 TGTGAGATGATCAAGAAC--AAGAGGAAAGAGGAAAGGAAAGATAGAGAGACT 602
Qy 181 LysProIleIleGlnThrArgArgProGluTyrThrProIleHisLeuSer 198
Db 603 AAGCCAAATTTATTCACAGACGAGCGGAGTACAGCGGATCCACTCAGC 656
RESULT 8
AK092184
LOCUS
DEFINITION
Homo sapiens CDNA FLJ34865 fis, clone NT2NE2014104, highly similar
to DOWN SYNDROME CRITICAL REGION PROTEIN 1.
ACCESSION
AK092184
VERSION
AK092184.1 GI:21750714
KEYWORDS
oligo capping; fis (full insert sequence).
SOURCE
Homo sapiens teratocarcinoma cell line:NT2 CDNA to mRNA,
clone lib:NT2NE2 clone:NT2NE2014104.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Nihomiyu, K., Magatsuma, M., Kanda, K., Kondo, H., Yokoi, T.,
Kodaishi, H., Furuya, T., Takahashi, M., Kikkawa, E., Omura, Y., Abe, K.,
Kamihara, K., Katsura, N., Sato, K., Tanikawa, M., Yamazaki, M.,
Yamamoto, J., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S.,
Suzumoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T.,
Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M.,
Kikuchi, H., Murakawa, K., Kanehori, K., Takahashi-Fujii, A.,
Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S.,
Nagahara, K., Masuno, Y., Nagai, K. and Isogai, T.
NEDO human CDNA sequencing project
Unpublished
2 (bases 1 to 2407)
Isogai, T. and Yamamoto, J.
Direct Submission
Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomic@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human CDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; CDNA full insert sequencing:
Research Association for Biotechnology (RAB); CDNA library
construction: Helix Research Institute (HRI) (supported by Japan

* Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.

FEATURES

source
1..2407
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NT2NE2014104"
/cell_lines="NT2"
/cell_type="teratocarcinoma"
/clone_lib="NT2NE2"
/note="Cloning vector: pME18SF3-3-mRNA from NT2 neuron after the differentiation of NT2 neuronal precursor cells."

BASE COUNT 607 a 519 c 627 g 654 t

ORIGIN

Alignment Scores:
Pred. No.: 9.67e-91 Length: 2407
Score: 1002.50 Matches: 191
Percent Similarity: 97.98% Conservative: 3
Best Local Similarity: 96.46% Mismatches: 3
Query Match: 95.66% Indels: 1
DB: 9 Gaps: 1

US-09-782-953-3 (1-198) x AK092184 (1-2407)

Qy 1 MetGluGluValAspLeuGlnAspLeuProSerAlaThrIleAlaCysHisLeuAspPro 20
Db 218 ATGGAGAGGTGGACCTCGACGACCTGCCAGCGCCACCATCGCTGTCACCTGGAGCCG 277
Qy 21 ArgValPheValAspGlyLeuCyArgAlaLysPheGluSerLeuPheArgThrTyxAsp 40
Db 278 CGCGTGTTCGTGGAGCGCTCGCGGCCAAATTTGAGTCCCTCTTAGACGTATGAC 337
Qy 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60
Db 338 AAGACATCACCCTTCAGTATTTTAAAGAGCTTCAACAGAGTCAGAAATAAACTTCAGCAAC 397
Qy 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
Db 398 CCTTCTCCGAGCAGATGCCAGCTCCAGCTGCATAGACTGAGTTCCTGGGAAGGAA 457
Qy 81 MetLysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProProAsn 100
Db 458 ATGAAGTTATATTTGCTCAGACCTTACATAGGAAGCTCACACCTGGCTCCGCCAAAT 517
Qy 101 ProAspLysGlnPheLeuLysSerProProAlaSerProProValGlyTyrLysGlnVal 120
Db 518 CCAGCAAGCAGTTTCTGATCTCCCTCCCGCTCTCCGCGAGTGGGATGGAAACAGTG 577
Qy 121 GluAspAlaThrProValIleAsnTyrAspLeuTyrAlaIleSerLysLeuGlyPro 140
Db 578 GAAGATCGACCCAGTCATAAACTATGATCTTATATGCCATCTCAAGCTTGGGGCCA 637
Qy 141 GlyGlyLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160
Db 638 GGGGAAAGATGATGAATTTGCAAGCGAGCTGACACCACTCCAGCGTGGTGGTCCATGTA 697
Qy 161 CysGluSerAspGlnGluAsnGluGluGluGluMetGluArgMetLysArgPro 180
Db 698 TGTGAGAGTGCATCAGAG---AAGAGGAAGAGAGGAATGAAAGATGAGAGACCT 754
Qy 181 LysProLysIleGlnThrArgArgProGluTyrThrProLysHisLeuSer 198
Db 755 AAGCAAAATATATCCAGACAGGAGCGGAGGTACAGCGCGATCCACCTCAGC 808

RESULT 9

AX365315
LOCUS AX365315
DEFINITION Sequence 5 from Patent WO0204491.
ACCESSION AX365315
linear PAT 15-FEB-2002

VERSION

AX365315.1 GI:18697045

KEYWORDS

house mouse.

SOURCE

Mus musculus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 Williams, S.R. and Rothermel, B.

Methods and compositions relating to muscle selective calcineurin

AUTHORS

interacting protein (mcip)

TITLE

Patent: WO 0204491-A 5 17-JAN-2002;

JOURNAL

Board of Regents, The University of Texas System (US); Williams,

FEATURES

Sanders R. (US); Rothermel, Beverly (US)

Location/Qualifiers

1..597

/organism="Mus musculus"

/db_xref="taxon:10090"

1..597

/note="unnamed protein product"

/codon_start=1

/protein_id="CAD23809.1"

/db_xref="GI:18697045"

/translation="MDFRDPSYFNFSLIACVANDVFSSESTRAKFESLFRTVKDDTT

FQYKSFKRVIRINFSNPLSADARLRLHKTEFLGKEMKLYFAQTLHGSSHLAPNPD

KQFLSPSPAPPVGWKQVEDATPVINDLYLALSKLGPGEKYLHATDPTFVVVHVH

CESDQNEEEEEEMERMPKPKIQTRRPEYPIHLS"

BASE COUNT 170 a 156 c 142 g 129 t

ORIGIN

Alignment Scores:
Pred. No.: 2.95e-83 Length: 597
Score: 920.50 Matches: 178
Percent Similarity: 92.42% Conservative: 5
Best Local Similarity: 89.90% Mismatches: 12
Query Match: 87.83% Indels: 3
DB: 6 Gaps: 1

US-09-782-953-3 (1-198) x AX365315 (1-597)

Qy 4 ValAspLeuGlnAspLeuPro-----SerAlaThrIleAlaCysHisLeuAspPro 20
Db 1 ATGGATTTTGGGACTTTAGCTACAAATTTAGCTCCCTGATTTGCTGTGGCAACGAT 60
Qy 21 ArgValPheValAspGlyLeuCyArgAlaLysPheGluSerLeuPheArgThrTyxAsp 40
Db 61 GATGCTTTCAGCGAAAGTGCAGACCGGCCAAATTTGAATTCCTCTCAGAACATATGAC 120
Qy 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60
Db 121 AAGGACACCACTTCCAGTATTTTAAAGACTTCAACAGTGTCCGGATATAAATTCAGCAAC 180
Qy 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
Db 181 CCTTATATTCGACGCGATGCCAGGCTGCGCTGCACAGACCGAGTTCCTGGGAAGGAA 240
Qy 81 MetLysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProProAsn 100
Db 241 ATGAAGTTGTATTTGCTCAGACTTTACATAGGAAGTTACACCTGGCTCCGCCCAAT 300
Qy 101 ProAspLysGlnPheLeuLysSerProProAlaSerProProValGlyTyrLysGlnVal 120
Db 301 CCCGCAAAACAGTCTCTCATCTCCCTCCGCGCTCTCTCCCTCCGCTGGTGGGAAACAAGTA 360
Qy 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro 140
Db 361 GAAGATGCCACCCCGTCATAAATTTACGATCTTTTATATGCCATCTCCAAGCTGGGGCA 420
Qy 141 GlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160
Db 421 GGAGAGAGTATGAAGTGCATGCAGCGACAGACCCACTCCAGTGTGGTGTCCACGTG 480
Qy 161 CysGluSerAspGlnGluAsnGluGluGluGluMetGluArgMetLysArgPro 180

Db	481	TGTGAGACTGACCAAGACAGATGAGAGGAAGACGAAGAGATGAGAGAAATGAGAGACCC	540
QY	181	LYSPProLySilelleGlnThrArgArgProGluTyrTThrProIleLeuSer	198
Db	541	AAGCCAAATCATCTCAGACACGAGACCGGAGTACACACCGATCCACTTAC	594
RESULT 10			
AF237790			
DEFINITION	Mus musculus myocyte-enriched calcineurin intercalin protein 1		
LOCUS	AF237790		
ACCESSION	AF237790		
VERSION	AF237790.1	GI:7542528	
KEYWORDS			
SOURCE	Mus musculus.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	1 (bases 1 to 597)		
TITLE	Rothenmel, B.A., Vega, R.B., Yang, J., Wu, H., Baszel-Dubay, R.S. and Williams, R.S.		
JOURNAL	A Protein Encoded within the Down Syndrome Critical Region is Enriched in Striated Muscles and Inhibits Calcineurin Signaling		
REFERENCE	J. Biol. Chem. (2000) In press		
AUTHORS	2 (bases 1 to 597)		
TITLE	Rothenmel, B.A., Vega, R.B., Yang, J., Wu, H., Baszel-Dubay, R.S. and Williams, R.S.		
JOURNAL	Direct Submission		
REFERENCE	Submitted (23-FEB-2000) Internal Medicine, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, TX 75390, USA		
FEATURES			
source	Location/Qualifiers		
	1..597		
	/organism="Mus musculus"		
	/db_xref="taxon:10090"		
	/chromosome="16"		
CDS	1..597		
	/note="similar to the EST with GenBank Accession Number AA200984; similar to Homo sapiens DSCR1"		
	/codon_start=1		
	/product="myocyte-enriched calcineurin intercalin protein		
	1 splice variant 4"		
	/protein_id="AA63486.1"		
	/db_xref="GI:7542529"		
	/translation="MDPRDPSYFSSLIACVNDVPSSETRAKESLFRTRYDKDTT		
	FOYFSPKRVIRINPSAADARLRLHTEPELGEMKLYFQGLHIGSSHLAPNPPI		
	KQFLSPASPVGKQVDAEDAPRVNLDLVLSLQPEKXELHAATDPFVSVVVH		
	CESDQENEEDEEMERKPKRKIIOTRPRETPPIHLS"		
BASE COUNT	170 a 156 c 142 g 129 t		
ORIGIN			
Alignment Scores:			
Prid. No.:	2,95e-83	Length:	597
Score:	920.50	Matches:	178
Percent Similarity:	92.42%	Conservative:	5
Best Local Similarity:	89.90%	Mismatches:	12
Query Match:	87.83%	Indels:	3
DB:	10	Gaps:	1
US-09-782-953-3 (1-198) x AF237790 (1-597)			
QY	4	ValAspLeuGlnAspLeuPro-----SerAlaThrIleAlaCySHisLeuAspPro	20
Db	1	ATGATTTTGGGACTTTAGCTACATTTTAAAGCTCCCTGATTCCTTGTGTGGCAAGAT	60
QY	21	ArgValIheValAspGlyLeuCyArgAlaLysPheGluSerLeuPheArgThrTyrAsp	40
Db	61	GATGTCTTCAGCGAAAGTGAAGACCGAGGCCCAATTGAATCCCTTCAGAACATAAGAC	120
QY	41	LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn	60
Db	121	AAGACACACACCTTCAGATATTTTAAAGAGCTTCAAAAGTGTCGGATAAACTTCAGCAAC	180

Oy		61	Proleuserr1A1AAAPRAATGTLauXLeuH1elYThrGluPheLauGIlytYGlu	80
Dd	181	CCTTATTCGGACGCAGATGCCAGGTCCGGTGCAACAAGCCGAATTCTCGGGAGAAGAA	240	
Oy	81	MethylserUyrPhealAGlnThrLeuH1elIeQYSerSerH1SleualAproProasn	100	
Dd	241	ATGAAGTTGTAATTTTGGCTCAGACTTTACACTAGTAGAAGTTCACACCCTGCTCCGCCCAT	300	
Oy	101	ProAbplyvGINPheLeu11eSerProProAlaserProProvalGIYTPrlyvGLInVal	120	
Dd	301	CCCCCAAAACGTTCCATCATCTCCCCTCCGCCCTCTCTCCCTGGCTGAAAACAAGTA	360	
Oy	121	GUABRAIAthrProVal11eAsnTYrAspLeuLeuTYrAl1eSerLYsLeuGIyPro	140	
Dd	361	GAAAGTGCACCCCCCGTCATAAATTAGCATTTTATATGCACTTCCAAGCTGGGCCA	420	
Oy	141	GIYGIULYTYRGILueH1eAla1a1aThrAspProThrProSerValValH1eVal	160	
Dd	421	GGAGAGAGATGAACACTGCATGCAGCCGACACCCCACTCCAGTGTGTGTCTCACGTG	480	
Oy	161	CysGISerArpGINguAsnGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu	180	
Dd	481	TGTGAGATGACCAAAGAAATAGAGAGAAAGAAAGAGATGAGAGATGAAAGAACCC	540	
Oy	181	LysProlYs11e1eGlnThrArgArgProGluTYrThrPro11eH1sLeuSer	198	
Dd	541	AAGCCCAAAATCATTCACACACGAGACCGAGATACACCGATCCACTTACG	594	
RESULT 11				
AF263239		2141 bp	mRNA	linear
LOCUS				ROD 30-MAY-2000
DEFINITION			Mus musculus calcineurin inhibitor mRNA, complete cds,	
ACCESSION		AF263239		
VERSION		AF263239.1	GI:8102011	
SOURCE				
KEYWORDS				
ORGANISM			Mus musculus.	
			Mus musculus	
			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
			Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
REFERENCE			1 (bases 1 to 2141)	
AUTHORS			Puentes,J.J., Pritchard,M.A., Pucharcos,C. and Estivill,X.	
TITLE			Down syndrome candidate region 1 (Dserr1), one of three	
JOURNAL			alternatively spliced exon 1 transcripts	
REFERENCE			unpublished	
AUTHORS			2 (bases 1 to 2141)	
TITLE			Puentes,J.J., Pritchard,M.A., Pucharcos,C. and Estivill,X.	
JOURNAL			Direct Submission	
			Submitted (02-MAY-2000) Centre de Genetica Medica i Molecular, IRO.	
			Avila, Casellasdefels Km. 2.7, L'Hospitalet de llobregat, Barcelona	
			08907, Spain	
FEATURES			Location/Qualifiers	
source			1..2141	
			/organism="Mus musculus"	
			/db_xref="taxon:10090"	
			/chromosome="16"	
CDS			31..627	
			/note="Down syndrome candidate region 1; Dserr1;	
			alternatively spliced exon 1"	
			/codon_start=1	
			/product="calcineurin inhibitor"	
			/protein_id="AAF72701.1"	
			/db_xref="gi:8102012"	
			/translation="MHPRDYSNFSLIACYANDVFSSETRAKFESLPFTYDKDTT	
			FQYPSFKVRINFNSPADARLRHKTFLEKEMCLYFAQTLIQSSHLAPNPPLA	
			KQFLSPASPPVGKVEDATPVNTYLVAISLGGEKXYELHAADTTPSVVVH	
			CESDDBNEEEDEMERMKRPKRKTQTRPRETPPHLS"	
BASE COUNT		505 a	527 c	558 g 548 t 3 others
ORIGIN				
Alignment Scores:				
Score: No.:		1 22e-81	length:	2141
Score:		911.00	Matches:	175

Percent Similarity: 94.68% Conservative: 3
 Best Local Similarity: 93.09% Mismatches: 10
 Query Match: 86.93% Indels: 0
 DB: 10 Gaps: 0

US-09-782-953-3 (1-198) x AF263239 (1-2141)

```

Qy 11 SerAlaThrIleAlaCysHisLeuAspProArgValPheValAspGlyLeuCysArgAla 30
Db 61 AGCTCCCTGATTGCTGTGCGCAACGATGATGCTTCAGCGAAGTGAGCAGGCGC 120
Qy 31 LysPheGluSerLeuPheArgThrTyrAspLysAspThrThrPheGlnTyrPheLysSer 50
Db 121 AAATTGATCCCTCTTCAGAACATATGACAGGACACCCCTTCAGATATTTAAAGC 180
Qy 51 PheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAspAlaArgLeuArg 70
Db 181 TTCAACGCTGTCGCGATAAATTCAGCAACCCCTTATCTGCGCGGATGCGAGCTCGG 240
Qy 71 LeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyrPheAlaGlnThrLeuHis 90
Db 241 CTGCACAGACCGGAGTTCCTGGGGAAGAAATGAAGTTGATTTTGTCTCAGACTTTAC 300
Qy 91 IleGlySerSerHisLeuAlaProAsnProAspLysGlnPheLeuIleSerProPro 110
Db 301 ATAGGAATTCACACTGGCTGCCCCCAATCCCGACAAACAGTTCCTCATCTCCCTCCG 360
Qy 111 AlaSerProValGlyTyrLysGlnValGluAspAlaThrProValIleAsnTyrAsp 130
Db 361 GCCTCTCTCCCTGCTGCGAACAAGTAGAAGATGCCACCCCGTCATATAATTAGCAT 420
Qy 131 LeuLeuTyrAlaIleSerLysLeuGlyProGlyLysTyrGluLeuHisAlaAlaThr 150
Db 421 CTTTTATGCAATCTCAAGCTGGGGCCAGGAGAGAAGATGAACATGTCAGCGACACA 480
Qy 151 AspProThrProSerValValHisValCysGluSerAspGlnGluAsnGluGlu 170
Db 481 GACACCACTCCAGTGTGTGTCACGTTGTGAGAGTACCACCAAGAGATGAGGAGAA 540
Qy 171 GluGluMetGluArgMetLysArgProLysProLysIleLeuGlnThrArgArgPro 190
Db 541 GAGGAAGAGATGGAGAGAATGAAGAGACCAAGCCCAAAATCATCCAGACAGGAGCG 600
Qy 191 GluTyrThrProIleHisLeuSer 198
Db 601 GAGTACACCATCCACCTCAGC 624

```

BC013551 2198 bp mRNA linear ROD 07-AUG-2002
 Mus musculus, Down syndrome critical region homolog 1 (human),
 clone MGC:19348 IMAGE:4236038, mRNA, complete cds.
 BC013551
 BC013551.1 GI:15488840
 MGC.
 house mouse.
 Mus musculus.
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 2198)
 Strausberg, R.
 Direct Submission
 Submitted (04-SEP-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LILNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome

Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>
 Contact: (Dickson, Mark) mcd@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 25 Row: m Column: 6.

FEATURES

```

source
1..2198
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="FVB/N"
/clone="MGC:19348 IMAGE:4236038"
/tissue_type="kidney, normal. 5 month old male mouse."
/clone_lib="NCI_CGAP_Kid14"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
77..673
/codon_start=1
/product="Down syndrome critical region homolog 1 (human)"
/protein_id="AAH13551.1"
/db_xref="GI:15488841"
/db_xref="LocusID:54720"
/translation="MHRDPSYNSFLIACVANDVDFSESETRAKFESLFRITYDKDTT
FOYKSKFRVRINFSNPISADARLRHLKTEFLKEMKLYFAQTUHLGSSHLAPNPDP
KQFLISPPASPPVGKQVEDATPVINYDLLTAISKLGPEKYTELHAATDTTSPVVVHV
CESDNEEBEEMERMKRPKIKIOTRPEYPIHLS"
BASE COUNT 533 a 539 c 569 g 557 t
ORIGIN

```

Alignment Scores:

```

Pred. No.: 1,268-81 Length: 2198
Score: 911.00 Matches: 175
Percent Similarity: 94.68% Conservative: 3
Best Local Similarity: 93.09% Mismatches: 10
Query Match: 86.93% Indels: 0
DB: 10 Gaps: 0

```

US-09-782-953-3 (1-198) x BC013551 (1-2198)

```

Qy 11 SerAlaThrIleAlaCysHisLeuAspProArgValPheValAspGlyLeuCysArgAla 30
Db 107 AGCTCCCTGATTGCTGTGCGCAACGATGATGCTTCAGCGAAGTGAGCAGGCGC 166
Qy 31 LysPheGluSerLeuPheArgThrTyrAspLysAspThrThrPheGlnTyrPheLysSer 50
Db 167 AAATTGATCCCTCTTCAGAACATATGACAGGACACCCCTTCAGTATTTAAAGC 226
Qy 51 PheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAspAlaArgLeuArg 70
Db 227 TTCAACGCTGTCGCGATAAATTCAGCAACCCCTTATCTGCGCGGATGCGAGCTCGG 286
Qy 71 LeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyrPheAlaGlnThrLeuHis 90
Db 287 CTGCACAGACCGGATCTCTGGGGAAGCAATGAAGTTGATTTTGTCTCAGACTTTAC 346
Qy 91 IleGlySerSerHisLeuAlaProAsnProAspLysGlnPheLeuIleSerProPro 110
Db 347 ATAGGAATTCACACTGGCTGCCCCCAATCCCGACAAACAGTTCCTCATCTCCCTCCG 406
Qy 111 AlaSerProValGlyTyrLysGlnValGluAspAlaThrProValIleAsnTyrAsp 130
Db 407 GCCTCTCTCCCTGCTGCGAACAAGTAGAAGATGCCACCCCGTCATATAATTAGCAT 466
Qy 131 LeuLeuTyrAlaIleSerLysLeuGlyProGlyLysTyrGluLeuHisAlaAlaThr 150
Db 467 CTTTTATGCAATCTCAAGCTGGGGCCAGGAGAGAAGTGAACATGTCAGCGACACA 526
Qy 151 AspProThrProSerValValHisValCysGluSerAspGlnGluAsnGluGlu 170
Db 527 GACACCACTCCAGTGTGTGTCACGTTGTGAGAGTACCACCAAGAGATGAGGAGAA 586

```

QY	171	GIUGUGUGMCTGIUARGMCTLYARPRPLYSPOLYGLIEGLIATHARGARPR	190
Db	587	GAGGAAGAGTGGAGAGTAAAGAGACCCCAAGCCCAATATCTCAGACGAGACG	646
QY	191	GIUTYRTHPRIOLEHLEUSER	198
Db	647	GAGTACACCCATCCACTCAGC	670
RESULT 13			
AF263240		2224 bp	linear
LOCUS			AF263240
DEFINITION		Mus musculus calcineurin inhibitor mRNA, complete cds,	ROD 30-MAY-2000
ACCESSION		AF263240	
VERSION		AF263240.1	GI:8102013
KEYWORDS			
SOURCE		Mus musculus.	
ORGANISM		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
REFERENCE		1 (bases 1 to 2224)	
AUTHORS		Fuentes,J.J., Pritchard,M.A., Pucharcos,C. and Estivill,X.	
TITLE		Down syndrome candidate region 1 (Dscr1), one of three alternatively spliced exon 1 transcripts	
JOURNAL		Unpublished	
REFERENCE		2 (bases 1 to 2224)	
AUTHORS		Fuentes,J.J., Pritchard,M.A., Pucharcos,C. and Estivill,X.	
TITLE		Direct Submission	
JOURNAL		Submitted (02-MAY-2000) Centre de Genetica Medica i Molecular, IRO, Avia, Casellesdefels Km. 2.7, L'Hospital del de Llobregat, Barcelona 08907, Spain	
FEATURES			
source		Location/Qualifiers	
		1..2224	
		/organism="Mus musculus"	
		/db_xref="taxon:10090"	
		/chromosome="16"	
CDS		354..710	
		/note="Down syndrome candidate region 1; Dscr1; alternatively spliced exon 1"	
		/codon_start=1	
		/product="calcineurin inhibitor"	
		/protein_id="AA072702.1"	
		/db_xref="GI:8102014"	
		/translation="MKLYPAQTLNIGSSHLAPNDPKFLISPPASPYGKQVEDATPVNLDLVAISKLGGEKXELHAATDTPSVVHVHVCSDQENEEBEMRMKPKKIOTRRPEYTPIHLS"	
BASE COUNT		519 a 554 c 566 g 562 t	3 others
ORIGIN			
Alignment Scores:			
Prod. No.:		1.61e-79	Length: 2224
Score:		890.00	Matches: 169
Percent Similarity:		99.41%	Conservative: 0
Best Local Similarity:		99.41%	Mismatches: 1
Query Match:		84.92%	Indels: 0
DB:		10	Gaps: 0
US-09-782-953-3 (1-198) x AF263240 (1-2224)			
QY	29	ATGATATLALYPHEGUSIRLEUPHEAARGTHYTRASPLYASAPHTRTHPREGINTYRPE	48
Db	198	AGGGCCAAATTTGAAATCCCTTCAGAACATATGACAAAGACACACCTTCAGATATTT	257
QY	49	LYSESRPHELYSARVALATGILAEASPHSERASNPROLEUSERALAAIAPALARY	68
Db	258	AAGAGCTTCAAACTGTGCCGATTAATTTCAAGACACCCCTTATCTGAGCGAGTCCAGG	317
QY	69	LEUARGLEUHSYRTHGILUPHELEUGILYSGIUMELYSEUTYRPEALAGINTHR	88
Db	318	CTGGCGCTGCACAGACGAGACTCTCTGGGGAAGAAATGAAGTTGATTTTGTCTCAGACT	377
QY	89	LEUHSIILEGYSERSRHSLAEUALAPROPROBANPPOABPLYSGLINPHEULEIUSER	108

[illegible]

Percent Similarity: 93.12% Conservative: 3
 Best Local Similarity: 91.53% Mismatches: 12
 Query Match: 84.69% Indels: 1
 DB: 10 Gaps: 1

US-09-782-953-3 (1-198) x AB075973 (1-626)

QY 11 SerAlaThrIleAlaCysHisLeuAspProArgValPheValAspGlyLeuCysArgAla 30
 DB 49 AGCTCCCTGATGCTGTGGCAACGGTGATGCTTTCAGCGAAGTGAACACCGGCC 108
 QY 31 LysPheGluSerLeuPheArgThrTyrAspLysAspThrThrPheGlnTyrPheLysSer 50
 DB 109 AAATTGATCCCTCTTCAGGACGTATGACAGACATCACCTTCCAGTATTTAAAGC 168
 QY 51 PheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAspAlaArgLeuArg 70
 DB 169 TTCAAGCGTGTCCGAATAAACTTCAGCAACCCCTGCTGCGCGGACGCCAGCTCGG 228
 QY 71 LeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyrPheAlaGlnThrLeuHis 90
 DB 229 CTGCATAGACAGAGTTCTTGGGGAAGGAGATGAAGCTGTACTTTGCCAGACTTTACAC 288
 QY 91 IleGlySerSerHisLeuAlaProAsnProAspLysGlnPheLeuLysSerProPro 110
 DB 289 ATAGGAAGTTACACCTGGCTGCCCAATCCAGACAAAGTTCATCTCTCCCT 348
 QY 111 AlaSerProValGlyTyrLysGlnValGluAspAlaThrProValIleAsnTyrAsp 130
 DB 349 GCCTCTCCACCTGTGGCTGGAAACAGTAGAGAGCTACCCAGTCATATAACTAGAT 408
 QY 131 LeuLeuTyrAlaIleSerLysLeuGlyProGlyGluLysTyrGluLeuHisAlaAlaThr 150
 DB 409 CTTTATATCTATCTCAAGCTGGGACCGAGAGAGTACGAGCTACACGCGACGA 468
 QY 151 AspProThrProSerValValHisValCysGluSerAspGln---GluAsnGluGlu 169
 DB 469 GACACACGCCCGCTGGTGTCCAGCTGTGTGAGAGTACCAAGAGGAGGAGGAA 528
 QY 170 GluGluGluMetGluArgMetLysArgProLysPheLysIleGlnThrArgArg 189
 DB 529 GAGGAGGAGGAGTGGAGAGATGAAGAGACCCCAAGCAAAATCATCCAGACTCGAGG 588
 QY 190 ProGluTyrThrProIleHisLeuSer 198
 DB 589 CCAGAGTACACCCATCCACCTCAGC 615

RESULT 15

CGU60263 2216 bp mRNA linear ROD 03-SEP-1997
 LOCUS Cricetulus griseus Adapt78 (adapt78) mRNA, complete cds.

DEFINITION U60263
 ACCESSION U60263.1 GI:2351390

VERSION
 KEYWORDS

SOURCE Cricetulus griseus.

ORGANISM

Cricetulus griseus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
 Cricetulus.

REFERENCE

1 (bases 1 to 2216)
 Crawford,D.R., Leahy,K.P., Abramova,N., Lan,L., Wang,Y. and
 Davies,K.J.
 TITLE Hamster adapt78 mRNA is a Down syndrome critical region homologue
 that is inducible by oxidative stress
 JOURNAL Arch. Biochem. Biophys. 342 (1), 6-12 (1997)
 MEDLINE 97329095
 PUBMED 9185608

REFERENCES

2 (bases 1 to 2216)
 Crawford,D.R., Leahy,K.L. and Davies,K.J.A.
 TITLE Direct Submission
 JOURNAL Submitted (10-JUN-1996) Biochem. and Mol. Biol., Albany Medical
 College, 47 New Scotland Avenue, Albany, NY 12208, USA

FEATURES

Location/Qualifiers

source

1..2216
 /organism="Cricetulus griseus"
 /db_xref="taxon:10029"
 /cell_type="HA-1"

gene

1..2216
 /gene="adapt78"

CDS

70..663
 /gene="adapt78"
 /note="similar to human DSCR1; an oxidant- and
 calcium-inducible mRNA"

/codon_start=1
 /product="Adapt78"

/protein_id="AAB68517.1"

/db_xref="GI:2351391"

/translation="MHFRDENVFSLIACVANGDVSESETRAKFESLFRYDKDIT
 FOYKSFYKVRINFENPISADARLQLHKTEFLGKEMKLYFACTLHIGSSHAPPND
 KOFLLSPASPVPVGMKQVEDATPIYINLLYLAISKLPGEKLYELHAATDTTPSVVVHV
 CSDQDENBEEEMERMKPKPKIITRYPEYTPHLS"

BASE COUNT 544 a 515 c 564 g 593 t

ORIGIN

Alignment Scores:

Pred. No.: 5,718-79 Length: 2216
 Score: 884.50 Matches: 172
 Percent Similarity: 93.62% Conservative: 4
 Best Local Similarity: 91.49% Mismatches: 11
 Query Match: 84.40% Indels: 1
 DB: 10 Gaps: 1

US-09-782-953-3 (1-198) x CGU60263 (1-2216)

QY 11 SerAlaThrIleAlaCysHisLeuAspProArgValPheValAspGlyLeuCysArgAla 30
 DB 100 AGCTCCCTGATGCTGTGGCAACGGTGATGCTTTCAGCGAAGTGAACACCGGCC 159
 QY 31 LysPheGluSerLeuPheArgThrTyrAspLysAspThrThrPheGlnTyrPheLysSer 50
 DB 160 AAATTGATCCCTCTTTCAGGACGTATGACAGGACATCACTTCCAGTATTTAAAGC 219
 QY 51 PheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAspAlaArgLeuArg 70
 DB 220 TTCAACGCTGTCCGAATAAACTTCAGCAACCCCTTATCCGAGCTACCCAGGCTCGAG 279
 QY 71 LeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyrPheAlaGlnThrLeuHis 90
 DB 280 CTGCATAGACAGGAGTTCCTGGGGAAGGAAATGAAGCTGTACTTTGCTCAGAGTTACAC 339
 QY 91 IleGlySerSerHisLeuAlaProAsnProAspLysGlnPheLeuLysSerProPro 110
 DB 340 ATAGGAAGCTCACCTGGCTCCGCCCAATCCAGACAAAGCAGTTCCTCATCTCTCTCCC 399
 QY 111 AlaSerProValGlyTyrLysGlnValGluAspAlaThrProValIleAsnTyrAsp 130
 DB 400 GCCTCTCCACCACTGTGTTGGAAAGCAAGTAGAAGACGCTACCCCAAGTATAAATACGAT 459
 QY 131 LeuLeuTyrAlaIleSerLysLeuGlyProGlyGluLysTyrGluLeuHisAlaAlaThr 150
 DB 460 CTTTATATGCTATCTCCAAGCTGGGCGCAGGTGAGAGTATGAGCTGCATCGACGACNA 519
 QY 151 AspProThrProSerValValHisValCysGluSerAspGlnGluAsnGluGluGlu 170
 DB 520 GACACCACTCCAGTGTGTGTAGTCCACGCTGTGTGAAGCGACCAAGAGAAT---GAAGAG 576
 QY 171 GluGluGluMetGluArgMetLysArgProLysPheLysIleGlnThrArgArgPro 190
 DB 577 GAGGAGGAGATGGAGAGATGAAGAGACCCCAAGGCAAAATTTATCCAGACGAGGCGCA 636
 QY 191 GluTyrThrProIleHisLeuSer 198
 DB 637 GAGTACACGCTTATCCACCTCAGC 660

Search completed: December 14, 2002, 21:31:58

Sun Dec 15 08:38:11 2002

us-09-782-953-3.rge

Job time : 1660.79 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 11:37:42 ; Search time 7.3457 Seconds
(without alignments)
1117.976 Million cell updates/sec

Title: US-09-782-953-3
Perfect score: 1048
Sequence: 1 MEEVDLQDLPSAIIACHLDP.....RPRKLIQTRRPTPIHLS 198

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1040	99.2	198	1	CCP1_MOUSE
2	1002.5	95.7	197	1	CCP1_HUMAN
3	884.5	84.4	197	1	CCP1_CRIGR
4	613.5	58.5	197	1	CCP2_MOUSE
5	597	57.0	239	1	CCP3_MOUSE
6	596.5	56.9	197	1	CCP2_HUMAN
7	596	56.9	241	1	CCP3_HUMAN
8	379.5	36.2	292	1	NLA_DROME
9	351	33.5	207	1	CCPL_CAEEL
10	95.5	9.1	163	1	YA9F_SCHPO
11	88	8.4	630	1	YCF2_OENVI
12	88	8.4	721	1	YCF2_OENPI
13	86.5	8.3	822	1	NAH1_CRIGR
14	86	8.2	338	1	GALE_MYCPN
15	83	7.9	561	1	SNTC_HUMAN
16	82.5	7.9	827	1	MOF_DROME
17	81.5	7.8	1790	1	SEPA_EMEHI
18	81	7.7	150	1	ALL5_HEVBR
19	80.5	7.7	280	1	UPPS_METJA
20	80.5	7.7	360	1	VTPX_TTV1
21	79.5	7.6	904	1	DLG1_HUMAN
22	79	7.5	1142	1	JAK1_HUMAN
23	78.5	7.5	211	1	RCN1_YEAST
24	78	7.4	467	1	INVO_MOUSE
25	77.5	7.4	627	1	DNAX_GUITH
26	77	7.3	1033	1	S190_YEAST
27	77	7.3	1912	1	CHD4_HUMAN
28	76.5	7.3	558	1	ORC2_XENLA
29	76	7.3	445	1	HH3R_MOUSE
30	76	7.3	445	1	HH3R_RAT
31	76	7.3	450	1	INVO_LEMCA
32	76	7.3	1460	1	PANI_YEAST
33	75.5	7.2	594	1	YKAS_YEAST

ALIGNMENTS

RESULT 1
CCP1_MOUSE
ID CCP1_MOUSE STANDARD; PRT; 198 AA.
AC Q9JHG6; Q9JKK3; Q9JKK2; Q9JK51; Q9JK50;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Calcipressin 1 (Down syndrome critical region protein 1 homolog)
DE (Myocyte-enriched calcineurin interacting protein 1) (MCIP1).
GN DSCR1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM A).
RC STRAIN=BALB/C; TISSUE=Brain;
RX MEDLINE=205334792; PubMed=11080588;
RA Scippoli P., Petrini M., Lenzi L., Carinci P., Zannotti M.;
RT "The murine DSCR1-like (Down syndrome candidate region 1) gene family:
RT conserved synteny with the human orthologous genes.";
RL Gene 257:223-232(2000).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS A AND B).
RX MEDLINE=20187590; PubMed=10722714;
RA Rotheimer B., Vega R.B., Yang J., Wu H., Bassel-Duby R.,
RA Williams R.S.;
RT "A protein encoded within the Down syndrome critical region is
RT enriched in striated muscles and inhibits calcineurin signaling.";
RL J. Biol. Chem. 275:8719-8725(2000).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM A), AND FUNCTION.
RC TISSUE=Fetal brain;
RX MEDLINE=21152920; PubMed=11231093;
RA Casas C., Martinez S., Pritchard M.A., Fuentes J.J., Nadal M.,
RA Guinera J., Arbones M., Florez J., Soriano E., Estivill X.,
RA Alcantara S.;
RT "Dscr1, a novel endogenous inhibitor of calcineurin signaling, is
RT expressed in the primitive ventricle of the heart and during
RT neurogenesis.";
RL Mech. Dev. 101:289-292(2001).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORMS B AND C).
RA Fuentes J.J., Pritchard M.A., Pucharcos C., Estivill X.;
RT "Down syndrome candidate region 1 (Dscr1), one of three alternatively
RT spliced exon 1 transcripts.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A. (ISOFORM C).
RC STRAIN=C57BL/6J; TISSUE=Embryonic stem cells;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Iwawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,

Q1186 caenorhabdi
P13834 oryctolagus
Q09293 caenorhabdi
Q92180 bos taurus
Q92974 homo sapien
Q42877 lycopersico
O88573 mus musculus
P40318 saccharomyc
O94966 homo sapien
P22105 homo sapien
P40366 saccharomyc
P39770 drosophila

34 75.5 7.2 621 1 YPD7 CAEEL
35 75.5 7.2 734 1 GYS1 RABIT
36 75 7.2 367 1 YQO1 CAEEL
37 75 7.2 427 1 G6NT BOVIN
38 75 7.2 893 1 GEPH HUMAN
39 75 7.2 1191 1 RPB2 LYCES
40 75 7.2 1217 1 AF4 MOUSE
41 75 7.2 1319 1 SSM4 YEAST
42 74.5 7.1 1371 1 UBRJ HUMAN
43 74.5 7.1 4289 1 TENX HUMAN
44 74 7.1 167 1 YJG5 YEAST
45 74 7.1 1355 1 SALM DROME

RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schirral L.M., Straubli F., Suzuki R., Tomita M., Wagner L., Mashio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Guenichev S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombasere P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Sessa T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wyshew-Boris A., Yoshida K., Haasegawa Y., Kawaji H., Kontseuki S.,
 RA Hayashizaki Y.,
 RA "Functional annotation of a full-length mouse cDNA collection";
 RT Nature 409:685-690(2001).
 CC -1- FUNCTION: Inhibits calcineurin-dependent transcriptional responses
 CC by binding to the catalytic domain of calcineurin A. Could play a
 CC role during central nervous system development.
 CC -1- ALTERNATIVE PRODUCTS: At least 3 isoforms; a/1 (shown here), b/4
 CC and c; are produced by alternative splicing.
 CC -1- TISSUE SPECIFICITY: Highly expressed in heart and skeletal muscle.
 CC Also expressed in all other tissues.
 CC -1- SIMILARITY: BELONGS TO THE DSCR1 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF282255; AAF91461.1; -
 CC EMBL: AF237789; AAF63485.1; -
 CC EMBL: AF237790; AAF63486.1; -
 CC EMBL: AF260717; AAF70343.1; -
 CC EMBL: AF263239; AAF72701.1; -
 CC EMBL: AF263240; AAF72702.1; -
 CC EMBL: AK010696; BAB27128.1; -
 CC MGD: MGI:1890564; Decr1.
 CC Alternative splicing.
 KW VARSPLIC 1 28 MEEVDLQDPSATICHLDPRVFVUNGLC -> MHRDRSYN
 FT VARSPLIC 1 80 MISSING (IN ISOFORM B).
 FT CONFLICT 152 152 T -> P (IN REF. 2).
 FT SEQUENCE 198 AA, 22706 MW, BA53340AF66F0582 CRC64;
 SQ
 Query Match 99.2%; Score 1040; DB 1; Length 198;
 Best Local Similarity 99.5%; Pred. No. 1,1e-83;
 Matches 197; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

DT 16-OCT-2001 (Rel. 40, last sequence update)
 DT 16-OCT-2001 (Rel. 40, last annotation update)
 DE Calcipressin 1 (Down syndrome critical region protein 1) (Myocyte-
 DE enriched calcineurin interacting protein 1) (MCIP1) (Adapt78).
 GN DSCR1 OR DSCR1 OR ADAPT78.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS 3 AND 4).
 RX MEDLINE=96121593; PubMed=8595418;
 RA Fuentes J.-U., Pritchard M.A., Planas A.M., Bosch A., Ferrer I.,
 RA Bativill X.,
 RT "A new human gene from the Down syndrome critical region encodes a
 RT proline-rich protein highly expressed in fetal brain and heart";
 RL Hum. Mol. Genet. 4:1935-1944(1995).
 RN [2]
 RP SEQUENCE FROM N.A. AND ALTERNATIVE SPLICING.
 RX MEDLINE=97468152; PubMed=9325060;
 RA Fuentes J.-U., Pritchard M.A., Bativill X.,
 RT "Genomic organization, alternative splicing, and expression patterns
 RT of the DSCR1 (Down syndrome candidate region 1) gene";
 RL Genomics 44:358-361(1997).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Lung;
 RA Strauberg R.,
 RT Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 1-164 FROM N.A. (ISOFORM 2).
 RC TISSUE=Breast;
 RA Crawford D.R., Leahy K.P., Davies K.J.A.,
 RT "Adapt78, a calcium and oxidant-inducible RNA";
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP CHARACTERIZATION.
 RX MEDLINE=20320698; PubMed=10861295;
 RA Puentes J.J., Genesca L., Kingsbury T.J., Cunningham K.W.,
 RA Perez-Riba M., Bativill X., de la Luna S.,
 RT "DSCR1, overexpressed in Down syndrome, is an inhibitor of
 RT calcineurin-mediated signaling pathways";
 RL Hum. Mol. Genet. 9:1681-1690(2000).
 CC -1- FUNCTION: Inhibits calcineurin-dependent transcriptional responses
 CC by binding to the catalytic domain of calcineurin A. Could play a
 CC role during central nervous system development.
 CC -1- ALTERNATIVE PRODUCTS: 4 ISOFORMS; 1 (SHOWN HERE), 2, 3 AND 4; ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: Highly expressed heart, brain and skeletal
 CC muscle. Also expressed in all other tissues.
 CC -1- INDUCTION: By calcium.
 CC -1- SIMILARITY: BELONGS TO THE DSCR1 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U28833; AAB81557.1; -
 CC EMBL: U85265; AAB84370.1; -
 CC EMBL: U85266; AAB84371.2; -
 CC EMBL: U85267; AAB84372.1; -
 CC EMBL: BC002864; AAH02864.1; -
 CC EMBL: U53821; AAF21218.1; -
 CC Genew; HGNC:3040; DSCR1.
 DR MIM: 602917; -
 DR Alternative splicing.
 KW DOMAIN 166 174 POLY-GLU.
 FT VARSPLIC 1 28 MEEVDLQDPSATICHLDPRVFVUNGLC -> MHRDRSYN
 FT FSSLIACVANDIFSESET (IN ISOFORM 2).

```
FT VARSPLIC 1 29 MEEVDQLDLPSTACHLDPRVFDGLCR -> MYV (IN
FT VARSPLIC 1 80 MISSING (IN ISOFORM 4).
FT CONFLICT 159 159 H -> Q (IN REF. 4).
SQ SEQUENCE 197 AA; 22637 MW; 1BD426BB88167E6C CRC64;

Query Match 95.7%; Score 1002.5; DB 1; Length 197;
Best Local Similarity 96.5%; Pred. No. 2e-80;
Matches 191; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

Qy 1 MEEVDQLDLPSTACHLDPRVFDGLCRKFESLRTYDKDTTFQYFKSKRVIRNFSN 60
Db 1 MEEVDQLDLPSTACHLDPRVFDGLCRKFESLRTYDKDTTFQYFKSKRVIRNFSN 60

Qy 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPPVGMKQV 120
Db 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPPVGMKQV 120

Qy 121 EDATPVINDLYLXSKLGPGEKVELHAATDTPSVVVHVHVCESQDNEEEMERMKRP 180
Db 121 EDATPVINDLYLXSKLGPGEKVELHAATDTPSVVVHVHVCESQDNEEEMERMKRP 180

Qy 181 KPKIQTTRRPEYTHLS 198
Db 180 KPKIQTTRRPEYTHLS 197

RESULT 3
CCP1 CRIGR STANDARD; PRT; 197 AA.
ID CCP1 CRIGR
AC O35847;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Calcipressin 1 (Down syndrome critical region protein 1 homolog)
DE (Oxidative-induced protein Adapt78).
GN DSCR1 OR ADAPT78.
OS Cricetus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetus.
OX NCBI_TaxID=10029;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97339095; PubMed=9185608;
RA Crawford D.R., Leahy K.P., Abramova N., Lan L., Wang Y., Davies K.J.;
RT "Hamster adapt78 mRNA is a Down syndrome critical region homologue
RT that is inducible by oxidative stress.";
RL Arch. Biochem. Biophys. 342:6-12(1997).
CC -!- FUNCTION: Inhibits calcineurin-dependent transcriptional responses
CC by binding to the catalytic domain of calcineurin A. Could play a
CC role during central nervous system development (By similarity).
CC -!- INDUCTION: OXIDANT-INDUCIBLE.
CC -!- SIMILARITY: BELONGS TO THE DSCR1 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U60263; AAB68517.1; -.
SQ SEQUENCE 197 AA; 22674 MW; 0EFC68FA55826439 CRC64;

Query Match 84.4%; Score 884.5; DB 1; Length 197;
Best Local Similarity 91.5%; Pred. No. 3.8e-70;
Matches 172; Conservative 4; Mismatches 11; Indels 1; Gaps 1;

Qy 11 SATIACHLDPRVFDGLCRKFESLRTYDKDTTFQYFKSKRVIRNFSNPLSAADARL 70
Db 11 SSLIACVANGDVSESTRAKFSFLTYDKDTTFQYFKSKRVIRNFSNPLSAADARLQ 70
```

```
Qy 71 LHKTEFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPPVGMKQVEDATPVINYD 130
Db 71 LHKTEFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPPVGMKQVEDATPVINYD 130

Qy 131 LLVAISKLGPEKVELHAATDTPSVVVHVHVCESQDNEEEMERMKRPKPIQTTRP 190
Db 131 LLVAISKLGPEKVELHAATDTPSVVVHVHVCESQDNEEEMERMKRPKPIQTTRP 189

Qy 191 EYTPIHLS 198
Db 190 EYTPIHLS 197

RESULT 4
CCP2 MOUSE STANDARD; PRT; 197 AA.
ID CCP2 MOUSE
AC Q9JHG2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Calcipressin 2 (Down syndrome candidate region 1-like protein 1)
DE (Myocyte-enriched calcineurin interacting protein 2) (MCIP2).
GN DSCR1L1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20187590; PubMed=10722714;
RA Rothermel B., Vega R.B., Yang J., Wu H., Bassel-Duby R.,
RA Williams R.S.;
RT "A protein encoded within the Down syndrome critical region is
RT enriched in striated muscles and inhibits calcineurin signaling.";
RL J. Biol. Chem. 275:8719-8725(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=BALB/C; TISSUE=Brain;
RX MEDLINE=20534792; PubMed=11080588;
RA Strioppoli P., Petrini M., Lenzi L., Carinci P., Zannotti M.;
RT "The murine DSCR1-like (Down syndrome candidate region 1) gene family:
RT conserved synteny with the human orthologous genes.";
RL Gene 257:223-232(2000).
CC -!- FUNCTION: Inhibits calcineurin-dependent transcriptional responses
CC by binding to the catalytic domain of calcineurin A. Could play a
CC role during central nervous system development.
CC -!- TISSUE SPECIFICITY: Highest expression in heart, skeletal muscle
CC and brain. Lower expression in all other tissues.
CC -!- SIMILARITY: BELONGS TO THE DSCR1 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF237791; AAF63487.1; -.
DR EMBL; AF237887; AAF62538.1; -.
DR MGD; MGI:1858219; Dscr1l1.
SQ SEQUENCE 197 AA; 22025 MW; CE1306B7B03E70F1 CRC64;

Query Match 58.5%; Score 613.5; DB 1; Length 197;
Best Local Similarity 63.3%; Pred. No. 1.5e-46;
Matches 119; Conservative 22; Mismatches 38; Indels 9; Gaps 2;

Qy 11 SATIACHLDPRVFDGLCRKFESLRTYDKDTTFQYFKSKRVIRNFSNPLSAADARL 70
Db 11 STLVCVNDVEVFTNQEVKEKEGLRTYDECVTLQFLFKSFRVRINFSPKSAARARIE 70

Qy 71 LHKTEFLGKEMKLYFAQTLHGSS-----HLAPPNDKQFLISPPASPPVGMKQVEDATPV 126
```

```

Db 71 LHEHQFRGKJLKYPAQVQTPETDGLHLAPPOAPKQFLSPSPSPVGMKPPISDAPV 130
QY 127 INYDLVIAISKLGPEKELHAATDPTPSVYVHVCESQDNEEBEEMERMRKPPKIIQ 186
Db 131 LNYDLVIAVAKLGPKEKELHAGTESTPSVYVHVCSDMEEBEDPK-----TSPKPKIIQ 185
QY 187 TRREPTPT 194
Db 186 TRRPLPFP 193

RESULT 5
CCP3_MOUSE STANDARD; PRT; 239 AA.
ID CCP3_MOUSE 09CX87;
16-OCT-2001 (Rel. 40, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Calcipressin 3 (Down syndrome candidate region 1-like protein 2)
DE (Myocyte-enriched calcineurin interacting protein 3) (MCIP3).
DS DSCR12.
CS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=BAJB/c; TISSUE=Brain;
RA Strittipoli P., Petrini M., Lenzi L., Carinci P., Zannotti M.;
RA "The murine DSCR1-like gene family."
RA Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RL [2]
RP SEQUENCE OF 40-239 FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryonic head;
RC MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batilov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochia H.,
RA Kuehl P., Lewis S., Matsumoto Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schmitt L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carinci P., de Bonaldo M.F.,
RA Brownstein M.O., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Guesting S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Kombeerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Saeki H., Sato K., Schenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kontuski S.,
RA Hayashizaki Y.;
RA "Functional annotation of a full-length mouse cDNA collection."
RA Nature 409:685-690(2001).
RL -1- FUNCTION: Inhibits calcineurin-dependent transcriptional responses
CC by binding to the catalytic domain of calcineurin A. Could play a
CC role during central nervous system development (By similarity).
CC -1- SIMILARITY: BELONGS TO THE DSCR1 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF237888; AAF62539.1; -
DR EMBL; AK019377; BAB31687.1; -
DR MGI; MGI:1658220; Dscr1l2.
SQ SEQUENCE 239 AA; 27153 MW; 1B2687B47BB4D272 CRC64;

```

```

Query Match 57.0%; Score 597; DB 1; Length 239;
Best Local Similarity 60.7%; Pred. No. 5, 1e-45;
Matches 122; Conservative 20; Mismatches 49; Indels 10; Gaps 2;

QY 2 EVDVLDLPSATLACHDPRVFDGLCRAPFESLFRYDQDTFOYFKSPKRVIRINSP 61
Db 38 EMDLSDLPSTSLFACSVHEAVFEVQEKERFALFTLYDOVTFOLFKSRRVRIRINSP 97
QY 62 LSAADARLRHKEFLKEMKLYFAQTLLIG---SSHLAPNPDKQFLISPPASPVGM 117
Db 98 EAAARARIELHSEFHRKJLKYPAQVQVSGEARDKSYLLPPTKQFLISPPASPVGM 157
QY 118 KQVEDATPVINYDLVIAISKLGPEKELHAATDPTPSVYVHVCESQDNEEBEEMER 177
Db 158 KQSDAMPVINYDLVIAISKLGPEKELHAGTESTPSVYVHVCSETEEBED-----T 211
QY 178 KRPKKIIOTRRREPTPT 198
Db 212 KNPQKIIOTRRREPTPTAUS 232

RESULT 6
CCP2_HUMAN STANDARD; PRT; 197 AA.
ID CCP2_HUMAN 014206;
30-MAY-2000 (Rel. 39, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
15-JUN-2002 (Rel. 41, Last annotation update)
Calcipressin 2 (Thyroid hormone-responsive protein ZAKI-4) (Down
DE syndrome candidate region 1-like 1) (Myocyte-enriched calcineurin
DE interacting protein 2) (MCIP2).
GN DSCR1L1 OR ZAKI4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fibroblast;
RX MEDLINE=96279929; PubMed=8662924;
RA Miyazaki T., Kanou Y., Murata Y., Ohmori S., Niwa T., Maeda K.,
RA Yamamura H., Seo H.;
RA "Molecular cloning of a novel thyroid hormone-responsive gene, ZAKI-4,
RT in human skin fibroblasts."
RL J. Biol. Chem. 271:14567-14571(1996).
CC -1- FUNCTION: Inhibits calcineurin-dependent transcriptional responses
CC by binding to the catalytic domain of calcineurin A. Could play a
CC role during central nervous system development.
CC -1- TISSUE SPECIFICITY: Expressed in fibroblasts, heart, brain, liver,
CC and skeletal muscle but not in placenta, lung, kidney and
CC pancreas.
CC -1- INDUCTION: By thyroid hormone.
CC -1- SIMILARITY: BELONGS TO THE DSCR1 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D83407; BAA11911.1; ALT_INIT.
DR Genew; HGNC:3041; DSCR1L1.
DR MIM; 604876; -
SQ SEQUENCE 197 AA; 21955 MW; AF39735F6661CBED CRC64;

Query Match 56.9%; Score 596.5; DB 1; Length 197;
Best Local Similarity 62.2%; Pred. No. 4, 4e-45;
Matches 117; Conservative 21; Mismatches 41; Indels 9; Gaps 2;

QY 11 SATIACHDPRVFDGLCRAPFESLFRYDQDTFOYFKSPKRVIRINSPISADALR 70

```


Wang Z.-Y., Wasmann D.A., Weinstock G.M., Weisenbach J., Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Yen R.-F., Zavert J.S., Zhan M., Zhang G., Zhao Q., Zheng L., Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 "The genome sequence of *Drosophila melanogaster*.";
 Science 287:2185-2195(2000).
 RT -1- FUNCTION: REQUIRED FOR ELONGATION OF MEIOSIS I SPINDLE.
 CC -1- SIMILARITY: BELONGS TO THE DSCR1 FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF147700; AAD33987.1; -
 CC EMBL: AF003712; AAF55285.1; -
 CC FLYBASE: FBgn0026629; nla.
 CC DR SEQUENCE 292 AA; 31423 MW; 64F1BPF56EAC6F9 CRC64;
 SQ
 Query Match 36.2%; Score 379.5; DB 1; Length 292;
 Best Local Similarity 40.7%; Pred. No. 5,5e-26;
 Matches 79; Conservative 42; Mismatches 60; Indels 13; Gaps 5;
 Oy 3 EVD---LQDLPSATTCALDPRVFDGLCRAPFESLFRFYDQDTTFQYKSKRVINPS 59
 Db 105 EVADASFDDLPTISIIYTNHSEVFANPELKHAMELFRFSSATFQWRSRRRLRVND 164
 Oy 60 NLSAADARLRLHKTPELKE-MKLYFAQTL-HIGSSHLAPPNDKFLISPPASPVGM 117
 Db 165 NMLAANAKIKHQEPNKKYVITCFAPQVPVPSKNIQOPAPVKQFLISPPASPVGM 224
 Oy 118 KOVEDATPVINDLYAISKLGPEKYEELHAATDPTPSVVHVCSQDNEEEBEMRM 177
 Db 225 EREBEPEPLVNDLALSLTPGSEHLPQSEDPALIVHTAML-----AETGPGLOV 279
 Oy 178 KRPKKIIOTRPE 191
 Db 280 KAP---IVQTKPE 290
 RESULT 9
 CCPL_CABEL STANDARD; PRT; 207 AA.
 ID PS3806; Q9U6V5;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Calcipressin-1-like protein (Down Syndrome candidate region 1-like protein).
 DE RCN-1 OR DSCR1L OR F54E7.7.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
 OC Rhabditidae; Poldoerinae; Caenorhabditis.
 NCBI_TaxID=6239;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=Briscot N2;
 RA Bentley D.;
 RL Submitted (JUN-1994) to the EMBL/GenBank/DBJ databases.
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN=Briscot N2;
 RA MEDLINE=2021370; PubMed=10756093;
 RA Striopol P., Lenz L., Petrini M., Carinci P., Zannotti M.;
 RA "A new gene family including DSCR1 (Down syndrome candidate region 1) and ZAK1-4: characterization from yeast to human and identification of DSCR1-like 2, a novel human member (DSCR1L2).";
 RT DSCR1-like 2, a novel human member (DSCR1L2).";
 RL Genomics 64:252-263(2000).
 CC -1- FUNCTION: Inhibits calcineurin-dependent transcriptional responses

by binding to the catalytic domain of calcineurin A (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE DSCR1 FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U00067; AAK20076.1; -
 CC EMBL: AF176115; AAF01683.1; -
 CC DR WormPep; F54E7.7; CE01318.
 CC FT CONFLICT 205 207
 CC G (IN REF. 1).
 CC SEQUENCE 207 AA; 23030 MW; 0154E308ABD5B79 CRC64;
 SQ
 Query Match 33.5%; Score 351; DB 1; Length 207;
 Best Local Similarity 42.0%; Pred. No. 1.1e-23;
 Matches 74; Conservative 34; Mismatches 64; Indels 4; Gaps 3;
 Oy 8 DLPSATTCALDPRVFDGLCRAPFESLFRFYDQDTTFQYKSKRVINPSNLADA 67
 Db 27 DLPNIIIVQPEDEVFNKDKANFSSLFQIEKHDFPLRSRRRVLFSSPENATA 86
 Oy 68 RLRLHKTPELKEMLKLYFAQTLH--GSSHLAPPNDKFLISPPASPVGMKQVEDATP 125
 Db 87 KLIVQSPFKHEKAFPAQRIYVANSQMLSPPLKQFLISPPSPVGMETKMP 146
 Oy 126 VI-NYDILYAIKLGPEKYEELHAATDPTPSVVHVCSQDNEEEBEMRMKRP 180
 Db 147 VVCNFDLMARLASPAIDKEYVHNGDELPAIIVPCETPID-VPSAITEMPRPRP 201
 RESULT 10
 YAPF_SCHPO STANDARD; PRT; 163 AA.
 ID YAPF_SCHPO
 AC Q09791;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical protein C13G6.15c in chromosome I.
 DE SPAC13G6.15C OR SPAC24B11.04C.
 GN Schizosaccharomyces pombe (Fission yeast).
 OS Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 CC Schizosaccharomyces.
 NCBI_TaxID=4896;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A., Sgouros K., Peat N., Hayles J., Baker S., Basham D., Bowman S., Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M., Collins M., Connor R., Cronin A., Davis P., Feltham J., Fraser A., Gentles S., Goble A., Hamlin N., Harris D., Hildrew J., Hodgson G., Holtroyd S., Hornby T., Howarth S., Huckle E.J., Hunt S., Jorgels K., James K., Jones L., Jones M., Leather S., McDonald S., McLean J., Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C., Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovich E., Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S., Skelton J., Simmonds M., Squares R., Squares S., Stevens K., Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S., Woodward J., Volkart G., Aert R., Robben J., Grympos B., Welter J., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Fritz C., Holzer E., Moerl D., Hilbert H., Borzym K., Langner I., Beck A., Lehnach R., Reinhardt R., Pohl T.M., Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B., Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Motier S., Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,

[illegible]

Db 158 LVHGLSDIVHG-LLEEGALVGSPTREEVEGTREEVEGTREEVEVEE

RESULT 13

NAH1_CRIGR STANDARD; PRT; 822 AA.

AC P48761.

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Sodium/hydrogen exchanger 1 (Na(+)/H(+) exchanger 1) (NHE-1).

GN SLC9A1 OR NHE1.

OS Cricetulus griseus (Chinese hamster).

OC Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae; Cricetulus.

OC NCBI_TaxID=10029;

OX NCBI_TaxID=10029;

RP SEQUENCE FROM N. A.

RA MEDLINE=93192332; PubMed=8383540;

RA Councilon L., Pouyssegur J.; Chinese hamster Na⁺/H⁺ exchanger NHE1. "Nucleotide sequence of the Chinese hamster Na⁺/H⁺ exchanger NHE1. " Biochim. Biophys. Acta 1172:343-345(1993).

RL FUNCTION: INVOLVED IN PH REGULATION TO ELIMINATE ACIDS GENERATED BY ACTIVE METABOLISM OR TO COUNTER ADVERSE ENVIRONMENTAL CONDITIONS. MAJOR PROTON EXTRUDING SYSTEM DRIVEN BY THE INWARD SODIUM ION CHEMICAL GRADIENT. PLAYS AN IMPORTANT ROLE IN SIGNAL TRANSDUCTION.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.

CC -1- PTM: PHOSPHORYLATED (POSSIBLE).

CC -1- SIMILARITY: BELONGS TO THE NA(+)/H(+) EXCHANGER FAMILY.

CC -1- CAUTION: THE NUMBER, LOCALIZATION AND DENOMINATION OF HYDROPHOBIC DOMAINS IN THE NA(+)/H(+) EXCHANGERS VARY AMONG AUTHORS.

CC -1- CAUTION: HYDROPHOBIC DOMAINS A, B AND L ARE NOT BELIEVED TO BE TRANSMEMBRANAL, BUT ONLY MEMBRANE-ASSOCIATED.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

CC -----

DR EMBL, X68970; CAA48771.1; -

DR InterPro: IPR000676; NaH_Exchange.

DR InterPro: IPR004709; NaH_exchange3.

DR Pfam: PF00999; Na_H_Exchange; 1.

DR PRINTS; PR01084; NAHEXCHNGR.

DR TIGRFAMs; TIGR00840; b_gpal; 1.

KM Transmembrane; Glycoprotein; Transport; Antiport; Sodium transport; Multigene family; Phosphorylation.

KW

FT DOMAIN 1 12

FT DOMAIN 13 32

FT DOMAIN 33 105

FT DOMAIN 106 127

FT DOMAIN 128 130

FT DOMAIN 131 150

FT DOMAIN 151 162

FT DOMAIN 163 183

FT DOMAIN 184 188

FT DOMAIN 189 210

FT DOMAIN 211 230

FT DOMAIN 231 251

FT DOMAIN 252 260

FT DOMAIN 261 282

FT DOMAIN 283 301

FT DOMAIN 302 322

FT DOMAIN 323 336

FT DOMAIN 337 357

FT DOMAIN 358 368

FT TRANSMEM 389 410

J (M8) (POTENTIAL).

FT DOMAIN 411 416

FT TRANSMEM 417 438

FT DOMAIN 439 452

FT DOMAIN 453 473

FT TRANSMEM 474 482

FT TRANSMEM 483 503

FT DOMAIN 504 822

FT CARBOHYD 374 374

SEQUENCE 822 AA; 92003 MW; E97C1ACD4EB8BDAA CRC64; (POTENTIAL).

Query Match 8.3%; Score 86.5; DB 1; Length 822;

Best Local Similarity 23.7%; Pred. No. 6.2; Indels 39; Gaps 8;

Matches 40; Conservative 26; Mismatches 64;

QY 44 TFQYKSPKFRINFSNPLSA--DARLRHKTPEFKEMKLYFAQTLH----- 90

Db 646 TRQRLRSYNNHTL-VADPYEAWNMQLRRQKARQEQKSNVLTVPARHLDSPTMSRR 704

QY 91 IGSSHLA-PNPDKQFLISPPASBPQKQVEDATPVINDLVKSLAGKREYELHAA 149

Db 705 IGSDPLAYEPKADLPVITIDPASP-----QSPSVDLVNEELAKV-----LGVN 749

QY 150 TDPFSSVAVHVCSDQENEEEMERKRPKPKI-----IQTRRPETP 194

Db 750 RDPF-----RLTRGEEDDEDEDEGVIMRKRPSSPQTDVFTAPMYSF 793

RESULT 14

GALE MYCPN STANDARD; PRT; 338 AA.

AC P75517;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE UDP-glucose 4-epimerase (EC 5.1.3.2) (Galactowaldenase) (UDP-galactose 4-epimerase).

GN GALE OR MPN257 OR MP576.

OS Mycoplasma pneumoniae.

OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.

OX NCBI_TaxID=2104;

RP SEQUENCE FROM N. A.

RC STRAIN=ATCC 29342 / M129;

RA MEDLINE=97105885; PubMed=8948633;

RA Himmelfreisch R., Hilbert H., Plagene H., Pirkl E., Li B.-C., Herrmann R.;

RT "Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae."

RL Nucleic Acids Res. 24:4420-4449(1996).

CC -1- CATALYTIC ACTIVITY: UDP-glucose = UDP-galactose.

CC -1- COFACTOR: NAD.

CC -1- PATHWAY: Galactose metabolism; third step.

CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE SUGAR EPIMERASE FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

CC -----

DR EMBL, AE000056; AAB96224.1; -

DR HSSP; P09147; IYK5.

DR InterPro: IPR001509; Epimerase_Dh.

DR Pfam: PF01370; Epimerase; 1.

DR TIGRFAMs; TIGR01179; gale; 1.

KM Isomerase; NAD; Galactose metabolism; Complete proteome.

FT NP_BIND 7 38

SEQUENCE 338 AA; 38132 MW; 9C50FP3856B8C03 CRC64;

Query Match 8.2%; Score 86; DB 1; Length 338;

THIS PAGE BLANK (user)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

QM protein - nucleic search, using frame_plus_p2n model

Run on: December 14, 2002, 18:11:26 ; Search time 138.531 Seconds
(without alignments)
3202.482 Million cell updates/sec

Title: US-09-782-953-9
Perfect score: 1041
Sequence: 1 MFAP5MDCVSTLVACVVDV.....SPKPIQTRRPGLPSPVSN 197

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-Q/cgn2_1/USPTO spool/US09782953/runat_11122002_114428_17330/app_query.fasta_1.1173
-DB=N Geneseq 101002 -OPMT=fastap -SUFFIX=rng -MINMATCH=0 -1-LOOPCL=0
-LOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09782953 @CGN_1_1_220 @runat_11122002_114428_17330 -NCPU=6 -ICPU=3
-NO XLPXY -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DSV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 101002.*
1: /SID2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1041	100.0	594	24	AAD30154	Mouse MCIP associa
2	1007	96.7	3184	24	AAD30157	Human MCIP associa
3	1007	96.7	3184	24	ABL61768	Colon adenocarcino
4	1007	96.7	3184	24	ABL65187	Lung cancer relate
5	910	87.4	1021	24	ABA91458	Rat Down syndrome
6	905	86.9	934	22	AAF25338	Nucleotide sequenc
7	905	86.9	3159	24	ABA91457	Human Down syndrom
8	658.5	63.3	828	24	AAD30158	Human MCIP associa
9	641	61.6	2331	24	AAD30155	Human MCIP associa
10	633.5	60.9	720	24	AAD30159	Human MCIP associa
11	632.5	60.8	597	24	AAD30153	Mouse MCIP associa
12	628	60.3	2348	24	AAS94805	Human DNA sequence
13	626.5	60.2	2212	24	AAD30156	Human MCIP associa
14	613.5	58.9	597	24	AAD30152	Human MCIP associa
15	601	57.7	2173	24	ABN96843	Gene #3341 used to
16	600	57.6	2174	20	AX01282	Human DSCR1 coding
17	595	57.2	615	24	ABA91463	Rat Down syndrome
18	584.5	56.1	2358	21	AAF18328	Lung cancer associ
19	564	54.2	599	24	AAD30151	Human MCIP associa
20	404	38.8	1820	23	ABL29301	Drosophila melanog
21	404	38.8	12550	23	ABL29300	Drosophila melanog
22	323	31.0	442	21	AC01774	Human secreted pro
23	294	28.2	531	24	ABA91461	Rat Down syndrome
24	223	21.4	412	22	ABA56172	Human foetal liver
25	223	21.4	412	22	ABA56172	Human foetal liver
26	223	21.4	412	22	AAK04364	Human brain expres
27	223	21.4	412	22	AAI14449	Probe #4382 for ge
28	223	21.4	412	22	AAI35821	Probe #4507 used t
29	223	21.4	412	22	AAI04272	Probe #4263 used t
30	223	21.4	412	24	ABSO4419	Human genome-deriv
31	223	21.4	446	22	ABA43406	Human breast cell
32	223	21.4	446	22	ABA53852	Human foetal liver
33	223	21.4	446	22	AAK023598	Probe #2064 for ge
34	223	21.4	446	22	AAK02113	Human brain expres
35	223	21.4	446	22	AAK27563	Human bone marrow
36	223	21.4	446	22	AAI12148	Probe #2081 for ge
37	223	21.4	446	22	AAI33493	Probe #2179 used t
38	223	21.4	446	22	AAI02061	Probe #2052 used t
39	169	16.2	446	24	ABSO2041	Human genome-deriv
40	169	16.2	486	22	ABA42132	Human breast cell
41	169	16.2	486	22	ABA52554	Human foetal liver
42	169	16.2	486	22	ABA22343	Probe #809 for gen
43	169	16.2	486	22	AAK00816	Human brain expres
44	169	16.2	486	22	AAK36270	Human bone marrow
45	169	16.2	486	22	AAI10903	Probe #836 for gen
46	169	16.2	486	22	AAI32163	Probe #849 used to

ALIGNMENTS

RESULT 1
AAD30154
ID AAD30154 standard; DNA; 594 BP.
XX
AC AAD30154;
XX
DT 17-MAY-2002 (first entry)
XX
DE Mouse MCIP associated DNA #4.
XX
KW Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;
KW heart failure; cardiomyopathy; heart disease; mouse; gens; ds.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
CDS 1..594
FT /*tag= a

/product= "Mouse MCIP associated protein #4"

MO200204491-A2.

17-JAN-2002.

06-JUL-2001; 2001WO-US21662.

07-JUL-2000; 2000US-216601P.

13-FEB-2001; 2001US-0782953.

(TEXA) UNIV TEXAS SYSTEM.

(WIL/) WILLIAMS S R.

(ROTH/) ROTHERMEL B.

Williams SR, Rothermel B;

WPI; 2002-179698/23.

P-PSDB; AAE18913.

Screening for modulators of muscle calcineurin interacting protein (MCIP) binding, expression or phosphorylation, useful for treating cardiac hypertrophy or heart failure, comprises mixing MCIP, calcineurin and a test compound -

Disclosure; Page 153-154; 174pp; English.

The invention relates to muscle calcineurin interacting proteins (MCIPs) and nucleic acid molecules encoding such proteins. MCIPs form a physical complex with the catalytic subunit of calcineurin and increased levels of MCIPs correspond to a reduced ability of calcineurin to stimulate transcription of certain target genes. The invention also relates to methods for identifying modulators of MCIP binding, expression or phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin may be used for treating cardiac hypertrophy and heart failure. Antibodies to MCIP can be used in characterizing the MCIP content of healthy and diseased tissues and subsequently for determining the presence or absence of cardiomyopathy or as predictor of heart disease. The present sequence is mouse MCIP associated DNA.

CC Note: This sequence has been described as mouse MCIP1 in the CC specification, however the sequence seems to be a polynucleotide CC encoding a MCIP associated protein.

SQ Sequence 594 BP; 148 A; 165 C; 149 G; 132 T; 0 other;

Alignment Scores:

Pred. No.:	9, 71e-99	Length:	594
Score:	1041.00	Matches:	197
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	24	Gaps:	0

US-09-782-953-9 (1-197) x AAD30154 (1-594)

QY 1 MetPcAlaProSerMetAcAspCyAspValSerThreValAlaCysValValaAspVal 20

DB 1 ATGCCAGCCCTGACGACGTGATGTTCCACTCTGCTCCCTGTGCGGATG 60

QY 21 GluValPheThrArgGlnGlnValGluValGluPheGlnGluLeuPheArgThrTyrAsp 40

DB 61 GAGGTCTTTCACATCAGAGGATTAGGAAATTCAGAGGACTGTTCCGACCTTATGAT 120

QY 41 GluCysValThrPheGlnLeuPheLysSerPheArgValArgIleAsnPheSerHis 60

DB 121 GAATGTGTACGCTTCAGCTTTAGAGTTCCAGCGGTTGGAATTAATTCAGGCAT 180

QY 61 ProLysSerAlaAlaArgAlaArgIleGluLeuHisGlnThrGlnPheArgGlyLys 80

DB 181 CCCAAATCTCAGACCCGTCGCCGATAGACTTCATAGACTCAGTTCAGAGGAAAGAG 240

QY 81 LeuLysLeuTyrPheAlaGlnValGlnThrProGluThrAspGlyAspLysLeuHisLeu 100

DB 241 CTAAACTCTACTTCGCCAGGTCCAGACCCAGACAGATGAGACAAACTGATTTG 300

QY 101 AlaProGlnProAlaValGlnPheLeuIleSerProProSerSerProValGly 120

DB 301 GCACCTCCACAGCCTGCCAAACAGTCTCATCTACCCCCCTTCATCTCTGTTGGC 360

QY 121 TrpLysProIleSerAspAlaThrProValLeuAsnTyrAspLeuTyrAlaValAla 140

DB 361 TGAAGCTTATCAGAGATGCCACACAGTCTCACTATGACTTCTTTAGCTGTGCC 420

QY 141 LysLeuGlyProGlyGluValTyrGluLeuHisAlaGlyThrGlnSerThrProSerVal 160

DB 421 AAACAGACCCAGAGAGAAATATGAGCTGACCTGGAATGATGATACCGAGGCTC 480

QY 161 ValValHisValCyAspSerAspMetGluGluGluGluAspProLysThrSerProLys 180

DB 481 GTGGTCATGTGTGACAGCAGACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540

QY 181 ProLysIleIleGlnThrArgArgProGlyLeuProProSerValSerAsn 197

DB 541 CCMAAAATCATTCAGACCCGCGGCTTCGCTTCACACCTTCGTGTCMAC 591

RESULT 2

AAD30157

ID AAD30157 standard; DNA; 3184 BP.

XX AAD30157;

AC AAD30157;

DT 17-MAY-2002 (first entry)

XX 17-MAY-2002 (first entry)

DE Human MCIP associated DNA #3.

XX Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;

KW heart failure; cardiomyopathy; heart disease; human; gene; ds.

XX Homo sapiens.

OS Homo sapiens.

FH Key Location/Qualifiers

FT CDS /tag= a

FT /product= "Human MCIP associated protein #3"

XX MO200204491-A2.

XX 17-JAN-2002.

PD 06-JUL-2001; 2001WO-US21662.

PF 07-JUL-2000; 2000US-216601P.

PR 13-FEB-2001; 2001US-0782953.

XX (TEXA) UNIV TEXAS SYSTEM.

PA (WIL/) WILLIAMS S R.

PA (ROTH/) ROTHERMEL B.

XX Williams SR, Rothermel B;

PI WPI; 2002-179698/23.

DR P-PSDB; AAE18913.

Screening for modulators of muscle calcineurin interacting protein (MCIP) binding, expression or phosphorylation, useful for treating cardiac hypertrophy or heart failure, comprises mixing MCIP, calcineurin and a test compound -

Disclosure; Page 163-165; 174pp; English.

The invention relates to muscle calcineurin interacting proteins (MCIPs) and nucleic acid molecules encoding such proteins. MCIPs form a physical complex with the catalytic subunit of calcineurin and increased levels of MCIPs correspond to a reduced ability of calcineurin to stimulate transcription of certain target genes. The invention also relates to methods for identifying modulators of MCIP binding, expression or

CC phosphorylation. Inhibitors or promoters of MCP binding to calcineurin
 CC may be used for treating cardiac hypertrophy and heart failure.
 CC Antibodies to MCP can be used in characterising the MCP content of
 CC healthy and diseased tissues and subsequently for determining the
 CC presence or absence of cardiomyopathy or as predictor of heart disease.
 CC The present sequence is human MCP associated DNA.
 CC Note: This sequence has been described as human MCP3 in the
 CC specification, however the sequence seems to be a polynucleotide encoding
 CC a MCP associated protein.

XX SQ Sequence 3184 BP; 921 A; 681 C; 657 G; 925 T; 0 other;

Alignment Scores:

Pred. No.: 2.88e-94 Length: 3184
 Score: 1007.00 Matches: 190
 Percent Similarity: 98.98% Conservatives: 5
 Best Local Similarity: 96.45% Mismatches: 2
 Query Match: 96.73% Indels: 0
 DB: 24 Gaps: 0

US-09-782-953-9 (1-197) x AAD30157 (1-3184)

Qy 1 MetProAlaProSerMetAspCysAspValSerThrLeuValAlaCysValValAspVal 20
 Db 190 ATGCCAGCCCTAGCATGGACTGTGATGTTTCCACTCTGGTTGCTGTGGTGGATGTC 249
 Qy 21 GluValPheThrAenGlnGluValLysGluLysPheGluGlyLeuPheArgThrTyrAsp 40
 Db 250 GAGGCTCTTACCAATCAGGAGGTTAAGGAAAAATTTGGGGGACTGTTTCGACTTATGAT 309
 Qy 41 GluCysValThrPheGlnLeuPheLysSerPheArgValAlaGlnPheAsnPheSerHis 60
 Db 310 GACTGTGTGACGTTCCAGCTATTTAAGAGTTTCAGACGTGTCCGTATAAACTTCAGCAAT 369
 Qy 61 ProllysSerAlaAlaArgAlaArgIleGluLeuHisGluThrGlnPheArgGlyLysLys 80
 Db 370 CCTAAATCTGCAGCCCGAGTAGATAGAGTTCATGAAACCCCAATTCAGAGGAAAAA 429
 Qy 81 LeuLysLeuTyrPheAlaGlnValGlnThrProGluThrAspGlyAspLysLeuHisLeu 100
 Db 430 TTAAGCTCTACTTTGCACAGGTTTCAGACTCCAGAGACAGATGAGCAAACTGCACCTG 489
 Qy 101 AlaProGlnProAlaLysGlnPheLysLeuSerProSerSerProProValGly 120
 Db 490 GCTCACCCAGCCGCTGCAACAGTTTCTATCTCGCCCTCTCTCCCTCCACCTGTTAGC 549
 Qy 121 TrpLysProIleSerAspAlaThrProValLeuAsnTyrAspLeuLeuTyrAlaValAla 140
 Db 550 TGGCAGCCCATCAACGATGCCAGCCAGTCTCAACTATGACCTCTCTATGCTGTGGCC 609
 Qy 141 LysLeuGlyProGlyGluLysTyrGluLeuHisAlaGlyThrGluSerThrProSerVal 160
 Db 610 AAACTAGGACGAGGAGAGATGAGTCCATGACGGGACTGAGTCCACCCCAAGTGC 669
 Qy 161 ValValHisValCysAspSerAspMetGluGluGluAspProLysThrSerProLys 180
 Db 670 GTCTGTGACGTGTGCGACAGTGCATAGAGGAAGAGGACCAAGACTTCCCCAAAG 729
 Qy 181 ProllysIleIleGlnThrArgArgProGlyLeuProProSerValSerAsn 197
 Db 730 CCAAAATATCCAAATCTCGCGCTCTGGCTGCCACCTCCGTTGTTCAAC 780

RESULT 3

ABL61768
 ID ABL61768 standard; DNA; 3184 BP.

XX ABL61768;

XX 15-MAY-2002 (first entry)

XX Colon adenocarcinoma related gene sequence SEQ ID NO:105.

XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;

KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
 KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
 KW gene; ds.
 XX Homo sapiens.
 OS WO200194629-A2.
 PN 13-DEC-2001.
 PD XX
 XX 30-MAY-2001; 2001WO-US10838.
 XX 05-JUN-2000; 2000US-209473P.
 PR 05-JUN-2000; 2000US-209531P.
 PR 18-SEP-2000; 2000US-233133P.
 PR 18-SEP-2000; 2000US-233617P.
 PR 20-SEP-2000; 2000US-234009P.
 PR 20-SEP-2000; 2000US-234034P.
 PR 20-SEP-2000; 2000US-234052P.
 PR 22-SEP-2000; 2000US-234509P.
 PR 22-SEP-2000; 2000US-234567P.
 PR 25-SEP-2000; 2000US-234923P.
 PR 25-SEP-2000; 2000US-234924P.
 PR 25-SEP-2000; 2000US-235077P.
 PR 25-SEP-2000; 2000US-235082P.
 PR 25-SEP-2000; 2000US-235134P.
 PR 25-SEP-2000; 2000US-235280P.
 PR 26-SEP-2000; 2000US-235637P.
 PR 26-SEP-2000; 2000US-235638P.
 PR 27-SEP-2000; 2000US-235711P.
 PR 27-SEP-2000; 2000US-235720P.
 PR 27-SEP-2000; 2000US-235840P.
 PR 27-SEP-2000; 2000US-235863P.
 PR 28-SEP-2000; 2000US-236028P.
 PR 28-SEP-2000; 2000US-236032P.
 PR 28-SEP-2000; 2000US-236033P.
 PR 28-SEP-2000; 2000US-236034P.
 PR 28-SEP-2000; 2000US-236109P.
 PR 28-SEP-2000; 2000US-236111P.
 PR 29-SEP-2000; 2000US-236842P.
 PR 29-SEP-2000; 2000US-236891P.
 PR 02-OCT-2000; 2000US-237172P.
 PR 02-OCT-2000; 2000US-237173P.
 PR 02-OCT-2000; 2000US-237278P.
 PR 02-OCT-2000; 2000US-237294P.
 PR 02-OCT-2000; 2000US-237295P.
 PR 02-OCT-2000; 2000US-237316P.
 PR 03-OCT-2000; 2000US-237425P.
 PR 03-OCT-2000; 2000US-237598P.
 PR 03-OCT-2000; 2000US-237604P.
 PR 03-OCT-2000; 2000US-237606P.
 PR 03-OCT-2000; 2000US-237608P.
 PR 01-NOV-2000; 2000US-244867P.
 PR 01-NOV-2000; 2000US-245084P.

(AVAL-) AVALON PHARM.

Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 Soppet DR, Weaver Z;

WPI; 2002-189264/24.

Screening for anti-neoplastic agent involves exposing cells to a
 chemical agent to be tested for anti-neoplastic activity, and
 determining a change in expression of a gene of a signature gene set -

Claim 1; SEQ ID 105; 44pp; English.

The present invention describes a method (M1) for screening for an
 anti-neoplastic agent. The method involves exposing cells to a chemical
 agent to be tested for anti-neoplastic activity, determining a change in
 expression of at least one gene (I) of a signature gene set, where (I)
 comprises a sequence (S) selected from 8447 sequences (given in ABL61664

CC to ABL70101), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilms' tumour.

CC
XX Sequence 3184 BP; 921 A; 681 C; 657 G; 925 T; 0 other;

Alignment Scores:

Pred. No.:	2 888-94	Length:	3184
Score:	1007.00	Matches:	190
Percent Similarity:	98.98%	Conservative:	5
Best Local Similarity:	96.45%	Mismatches:	2
Query Match:	96.73%	Indels:	0
DB:	24	Gaps:	0

US-09-782-953-9 (1-197) x ABL61768 (1-3184)

OY 1 MetProAlAProSerMetAProCyAePValSerThrLeuValAlaCyAeValAaPVal 20
DB 190 ATGCCAGCCCTTACAGTATGATGCTGATGTTTCCACTCTGGTCCCTGTGGTGATGTC 249
OY 21 GluValAlaPheThrAngInGluValLyGluValPheGluGlyLeuPheArgThrTyraP 40
DB 250 GAGGTCTTTACCATAGAGGGTTAAGAAATTGGGGGACTGTTGGACTTATGAT 309
OY 41 GluCyValAlaThrPheGluLeuPheLeuSerPheArgValAlaArgIleAsnPheSerHis 60
DB 310 GACTGTGACGCTTCCAGCTATTATAGAGTTTACAGGTGTCCTGCTTAACTTCAGCAAT 369
OY 61 ProLySerAlaAlaArgAlaArgIleGluLeuHisGluThrGlnPheArgIleGlyLys 80
DB 370 CCGAATTCGACGCCGAGCTAGATAGAGCTTCATGAATCCCAATTCAGAGGAAAAA 429
OY 81 LeuValLeuValPheAlaGlnValGlnThrProGluThrAspGlyAspLysLeuHisLeu 100
DB 430 TTAAACCTCTACTTTCACAGGTTCAGACTCCAGACACAGATGAAACAACTGACCTTG 489
OY 101 AlaProGlnProAlaValysGlnPheLeuIleSerProProSerSerProProValGly 120
DB 490 GCTCCACCCGAGCTGCCAAACAGTTTCTCATCTCCGCCCTTCTCCCACTGTAGC 549
OY 121 TrpLyProIleSerAspAlaThrProValLeuAsnTyraSpLeuValAlaAla 140
DB 550 TGGCAGCCCATCAACATGCAAGTCCAGCCAGTCTCACTATGACCTCTATATGCTGGGCC 609
OY 141 LysLeuGlyProGlyGlyLysTyrgLysLeuHisAlaGlyThrGluSerThrProSerVal 160
DB 610 AAACCTAGACACAGAGAGAAAGTATGAGCTCCATGACGAGGACTAGTCCACCCCAATGTC 669
OY 161 ValValHisValAlaValAspSerAspMetGluGluGluAspProLysThrSerProLys 180
DB 670 GTGTCACACCTGTGCACAGTGCATGACAGAAAGAAAGAGAACCCAAAGATTTCCCAAG 729
OY 181 ProLysIleIleGlnThrArgArgProGlyLeuProProSerValSerAsn 197
DB 730 CCABAAATCTCCAAACTGGCGCTCCGCTCCCAACCCCTCGTGCAC 780

RESULT 4
ABL65187
ID ABL65187 standard; DNA; 3184 BP.

AC ABL65187;
XX
DT 15-MAY-2002 (first entry)
XX

DE Lung cancer related gene sequence SEQ ID NO:3524.

XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;
KW gene; ds.

OS Homo sapiens.

PN MO200194629-A2.

PD 13-DEC-2001.

PF 30-MAY-2001; 2001KO-US10838.

XX
XX 05-JUN-2000; 2000US-209473P.
PR 05-JUN-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-231133P.
PR 18-SEP-2000; 2000US-231617P.
PR 20-SEP-2000; 2000US-234039P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.

XX (AVALON) AVALON PHARM.

XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;

XX Soppe DR, Weaver Z;

XX WPI; 2002-188264/24.

PT Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set
PS Claim 1; SEQ ID 3524; 44pp; English.

CC The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical

CC agent to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (1) of a signature gene set, where (1)
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (1) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening
 CC an anti-neoplastic agent, and can be used for producing a product which
 CC is the data collected with respect to the anti-neoplastic agent as a
 CC result of M1, and the data is sufficient to convey the chemical
 CC structure and/or properties of the agent. M1 can be used in the
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilms' tumor.

SQ Sequence 3184 BP; 921 A; 681 C; 657 G; 925 T; 0 other;

Alignment Scores:
 Pred. No.: 2.88e-94 Length: 3184
 Score: 1007.00 Matches: 190
 Percent Similarity: 98.98% Conservative: 5
 Best Local Similarity: 96.45% Mismatches: 2
 Query Match: 96.73% Indels: 0
 DB: 24 Gaps: 0

US-09-782-953-9 (1-197) x ABL65187 (1-3184)

Qy 1 MetProAlaProSerMetAspCysAspValSerThrLeuValAlaCysValValAspVal 20
 Db 190 ATGCCAGCCCTAGCATGGAGCTGTGATGTTTCCACTCTGGTGTGCTGTGGATGTC 249
 Qy 21 GluValPheThrAenGlnGluValLysGluLysPheGluLysLeuPheArgThrTyrAsp 40
 Db 250 GAGGCTCTTTACCAATCAGGAGGTTAAGAAATTTGGGGAGCTGTTTCGACTTATGAT 309
 Qy 41 GluCysValThrPheGlnLeuPheLysSerPheArgValArgIleAenPheSerHis 60
 Db 310 GACTGTGTGAGCTTCCAGCTATTATTAAGAGTTTCAGAGCTGTCCGTATATAACTTCAGCAAT 369
 Qy 61 ProLysSerAlaAlaArgAlaArgIleGluLeuHisGluThrGlnPheArgGlyLysLys 80
 Db 370 CCTAAATCTGCAGCCCGAGCTAGATAGAGCTTCATGAAACCCCAATTCAGAGGAAATAA 429
 Qy 81 LeuLysLeuTyrPheAlaGlnValGlnThrProGluThrAspGlyAspLysLeuHisLeu 100
 Db 430 TTAAGCTCTACTTTGCACAGGTTTCAGACTCCAGAGACAGATGGAGCAAACTGCACCTG 489
 Qy 101 AlaProGlnProAlaLysGlnPheLeuLysSerProProSerSerProProValGly 120
 Db 490 GCTCCACCCAGGCTGCCAAGAGTTTCTCATCTCGCCCTTCTCTCCACCTGTAGC 549
 Qy 121 TrpLysProLysSerAspAlaThrProValLeuAsnTyrAspLeuLeuTyrAlaValAla 140
 Db 550 TGGCAGCCCATCAACGATGCCAGCCAGTCTCACTATGATGCTGTGGCC 609
 Qy 141 LysLeuGlyProGlyGluLysTyrGluLeuHisAlaGlyThrGluSerThrProSerVal 160
 Db 610 AAATAAGGACCCAGAGAGAGATGATGAGCTCCATGCGGGAGCTGAGTCCACCCCAAGTGC 669
 Qy 161 ValValHisValCysAspSerAspMetGluGluGluAspProLysThrSerProLys 180
 Db 670 GTGTGACCTGTGCGACAGTGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 729
 Qy 181 ProLysIleLeuGlnThrArgArgProGlyLeuProProSerValSerAsn 197
 Db 730 CCAAAATATCAAACTCGCGCTGCTGGCTCCACCTCCCTGCTGCCAAC 780

RESULT 5

ABA91458/c

ID ABA91458 standard; cDNA; 1021 BP.

XX

AC ABA91458;

XX 18-APR-2002 (first entry)
 XX Rat Down syndrome critical region 1-like 1 protein related clone.
 DE XX
 XX Down syndrome critical region 1-like 1; DSCR1L1 alpha; rat;
 KW Down syndrome; Alzheimer's disease; dementia; transgenic animal;
 KW animal model; diagnosis; gene therapy; ss.
 XX OS
 XX Rattus norvegicus.
 XX PN
 XX WO200204513-A2.
 XX 17-JAN-2002.
 XX
 PF 11-JUL-2001; 2001WO-US21982.
 XX
 PR 11-JUL-2000; 2000US-0614474.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Loring JF, Tingley DW, Edwards CM, Streeter DG;
 XX
 DR WPI; 2002-164633/21.
 XX
 PT Novel Down syndrome critical region 1-like 1 protein and nucleic acid
 PT encoding the protein useful for diagnosis and treatment of Alzheimer's
 PT disease, Down syndrome and other forms of dementia -
 PS Claim 3; Page 48; 54pp; English.
 XX
 CC The present sequence is that of cDNA clone Incyte ID No: 219435-1,
 CC which shows 88% identity to nucleotides 443-954 of a cDNA clone
 CC encoding human Down syndrome critical region 1-like 1 protein
 CC (DSCR1L1 alpha, see ABA91457), and 87% identity to nucleotides
 CC 1040-1212 of the human sequence. The nucleic acid molecule is
 CC useful for producing transgenic cell lines or organisms which model
 CC human disorders and upon which potential therapeutic treatments
 CC for such disorders may be tested. Differential expression of the
 CC human DSCR1L1 alpha gene is diagnostic of Down syndrome, Alzheimer's
 CC disease and other forms of dementia.
 XX
 SQ Sequence 1021 BP; 237 A; 241 C; 286 G; 257 T; 0 other;

Alignment Scores:
 Pred. No.: 7.39e-85 Length: 1021
 Score: 910.00 Matches: 192
 Percent Similarity: 94.63% Conservative: 2
 Best Local Similarity: 93.66% Mismatches: 3
 Query Match: 87.42% Indels: 8
 DB: 24 Gaps: 0

US-09-782-953-9 (1-197) x ABA91458 (1-1021)

Qy 1 MetProAlaProSerMetAspCysAspValSerThrLeuValAlaCysValValAspVal 20
 Db 855 ATGCCAGCCCTAGCATGGAGCTGTGATGTTTCCACTCTGGTCCCTGTGTGGTGTG 796
 Qy 21 GluValPheThrAsnGlnGluValLysGluLysPheGluGlyLeuPheArgThrTyrAsp 40
 Db 795 GAGGCTCTTTACCAATCAGGAGGTTAAGAAAAATTTGAGGAGCTGTTCCGGACCTATGAC 736
 Qy 41 GluCysValThrPheGlnLeuPheLysSerPheArgArgValArg-IleAsnPheSerHis 60
 Db 735 GAAATGTGACCTTTTCAGCTGTTTAAGAGTTTCCAGCCGGTTCCGAATAAATTTTCAGCCA 676
 Qy 60 sProLysSerAlaAlaArg-AlaArgIle-GluLeuHisGlu-ThrGlnPheArg-GlyL 79
 Db 675 CCCCCAAGCTGCAGCCCGTGGCCGATAGGAGCTTCATGTGGACCCCATGTCAGGGA 616
 Qy 79 ysLysLeuLysLeuTyrPheAlaGlnValGlnThrProGluThrAspGly-AspLysLeu 98
 Db 615 AGAAGCTGAAACTCTACTTTTCACAGGTTCCAGAGTCCAGACCCAGAGACAGATGGTAGACAACTG 556

QY 99 His-LeuAlaProGlnProAlaIysGlnPheLeuIleSerProProSerSerProPr 118
 DB 555 CAGTTGGACACCCCAACCCCTGCGCAACAGTTCTCTCATCTACCCCTTATCTCTCC 496
 QY 118 oValGlyTrpLysProIleSerAspAlaThrProValLeuAen-TyrAspLeuLeuTyrA 138
 DB 495 CGTTGGCTGGAGCCATCAGCGATGCCACACAGTCCCAAGTACGACCTCTTATG 436
 QY 138 lAvalAlaIysLeuGlyProGlyGlyLysTyrGlnLeuHisAlaGlyThrGlySerThrP 158
 DB 435 CCGTGGCCAACTAGACACGAGAGAAATATAGCTGATCGGGAAGCTGATCTACAC 376
 QY 158 rGSerValAlaValHisValCysAspSerAspMetGluGluGluLysProLysThrS 178
 DB 375 CCAAGCTTCTGTCACGCTGTGTGACAGGACTTGGAGAGAGAGAGATCCAAAGACTT 316
 QY 178 ePrLysPProLysIleIleGlnThrArgArgProGlyLeuProProSerValSerAen 197
 DB 315 CCCCAAGCCAAATCATCTCAGACCCGGCGTCTGCTGCTGCTGCTGCTGCTCAAC 257
 RESULT 6
 AAF25338
 ID AAF25338 standard; cDNA; 934 BP.
 AC AAF25338;
 XX
 DT 30-APR-2001 (first entry)
 DE Nucleotide sequence of a human detoxification protein.
 XX
 KM Human; detoxification protein; DEXT; cancer; leukemia; melanoma;
 KM adenocarcinoma; autoimmune disorder; inflammatory disorder;
 KM rheumatoid arthritis; asthma; atherosclerosis; autoimmune thyroiditis;
 KM psoriasis; ulcerative colitis; infection; cell proliferative disorder;
 KM actinic keratosis; arteriosclerosis; cirrhosis; hepatitis; ss.
 OS Homo sapiens.
 XX
 PH Key location/Qualifiers
 FT CDS 20..724
 FT /tag= a
 FT /product= "detoxification protein"
 FT sig_peptide 20..100
 FT /tag= b
 XX
 PN WO200104305-A2.
 PD 18-JAN-2001.
 PF 06-JUL-2000; 2000WO-US18509.
 PR 07-JUL-1999; 99US-0142678.
 PA (INCY-) INCYTE GENOMICS INC.
 PI Tang YT, Yue H;
 DR MPI; 2001-147293/15.
 DR P-PSDB; AAB31788.
 XX
 PT New human detoxification protein and polynucleotide, useful for
 PT diagnosis, prevention and treatment of autoimmune/inflammatory
 PT disorders and cell proliferative disorders including cancer -
 XX
 PS Claim 5; Page 79; 79pp; English.
 XX
 CC The present sequence encodes a human detoxification polypeptide (DETX).
 CC DETX and its (ant)agonists are useful for preventing or treating
 CC disorders associated with decreased or increased expression or activity
 CC of DETX. DETX polypeptides are useful for screening compounds that
 CC specifically binds to DETX and for identifying (ant)agonists.
 CC Diseases prevented, treated and diagnosed include cancers (e.g.

CC leukemia, melanoma, adenocarcinoma, cancers of the brain, bone marrow,
 CC breast, kidney, liver, pancreas, prostate and uterus),
 CC autoimmune/inflammatory disorders (e.g. rheumatoid arthritis, asthma,
 CC atherosclerosis, autoimmune thyroiditis, psoriasis and ulcerative
 CC colitis), bacterial, fungal, parasitic infections and cell
 CC proliferative disorders (e.g. actinic keratosis, arteriosclerosis,
 CC cirrhosis and hepatitis). Anti-DETX antibodies may be used as
 CC antagonists, as a targeting or delivery mechanism for bringing
 CC pharmaceutical agents into contact with cells or tissues expressing
 CC DETX and for diagnosis of DETX-related disorders.
 XX
 SO Sequence 934 BP; 247 A; 244 C; 206 G; 237 T; 0 other;
 Alignment Scores:
 Pred. No.: 2.17e-84 Length: 934
 Score: 905.00 Matches: 171
 Percent Similarity: 95.72% Conservative: 8
 Best Local Similarity: 91.44% Mismatches: 8
 Query Match: 86.94% Indels: 0
 DB: 22 Gaps: 0
 US-09-782-953-9 (1-197) x AAF25338 (1-934)
 QY 11 SerThrLeuValAlaCysValAlaAspValGluValPheThrAngGlnValIysGlu 30
 DB 161 AACTGTTGTTGGTGTGCAATGTTCCACGTCAGTGTGTTGAGAGAGAGAGAGAGAG 220
 QY 31 LysPheGluGlyLeuPheArgThrTyrAspGluCysValThrPheGlnLeuPheLysSer 50
 DB 221 AAATTTAGGAGACTGTTTCCGACTTATGATGACTGTGTGAGCTTCCAGCTATTAAAGT 280
 QY 51 PheArgArgValAlaGlyLeuAsnPheSerHisProLysSerAlaAlaArgAlaArgIleGlu 70
 DB 281 TTCAGACCTGTCGGATATTAACCTTCAGCAATCTTAATTCAGCCGAGCTAGAGATAGAG 340
 QY 71 LeuHisGlnThrGlnPheArgGlyLysLysLeuLysLeuTyrPheAlaGlnValAlaGlnThr 90
 DB 341 CTTCAAGAAACCAATTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 400
 QY 91 ProGlnThrAspGlyAspLysLeuHisLeuAlaProProGlnProAlaIysGlnPheLeu 110
 DB 401 CCAGAGACAGATGAG 460
 QY 111 IleSerProProSerSerProProValGlyTrpLysProIleSerAspAlaThrProVal 130
 DB 461 ATCTGCCCCCTTCTCTCCCACTGTGTGCTGCGCCATCAAGATGACAGCCCACTG 520
 QY 131 LeuAsnTyrAspLeuLeuTyrAlaValAlaIysLeuGlyProGlyGlyLysTyrGlnLeu 150
 DB 521 CTCACCTATGACTCTCTCTATGCTGTGCGCAACTAGAGACAGAGAGAGAGATAGAGTC 580
 QY 151 HisAlaGlyThrGlnSerThrProSerValAlaValHisValCysAspSerAspMetGlu 170
 DB 581 CATGCAGGAGACTGAGTCCACCCCAAGTGTGCGTGCAGAGTGCAGAGTGCATAGAG 640
 QY 171 GlnGluGluAspProLysThrSerProLysProLysIleIleGlnThrArgArgProGly 190
 DB 641 GAAG 700
 QY 191 LeuProProSerValSerAen 197
 DB 701 CTGCCACCTCCGTGTCCAC 721
 RESULT 7
 ID ABA91457
 ID ABA91457 standard; cDNA; 3159 BP.
 AC ABA91457;
 XX
 DT 18-APR-2002 (first entry)
 XX
 DE Human Down syndrome critical region 1-like 1 protein cDNA.

PT Screening for modulators of muscle calcineurin interacting protein
 PT (MCIP) binding, expression or phosphorylation, useful for treating
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,
 PT calcineurin and a test compound -

XX Example 1; Page 167-168; 174pp; English.

XX The invention relates to muscle calcineurin interacting proteins (MCIPs)
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical
 CC complex with the catalytic subunit of calcineurin and increased levels
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate
 CC transcription of certain target genes. The invention also relates to
 CC methods for identifying modulators of MCIP binding, expression or
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin
 CC may be used for treating cardiac hypertrophy and heart failure.
 CC Antibodies to MCIP can be used in characterizing the MCIP content of
 CC healthy and diseased tissues and subsequently for determining the
 CC presence or absence of cardiomyopathy or as predictor of heart disease.
 CC The present sequence is human MCIP associated DNA.
 CC Note: This sequence has been described as splice variant of MCIP1
 CC initiated by exon 4 in the specification, however the sequence seems
 CC to be a polynucleotide encoding a MCIP associated protein.

XX Sequence 828 BP; 227 A; 188 C; 227 G; 186 T; 0 other;

XX Alignment Scores:

Pred. No.:	6,52e-59	Length:	828
Score:	658.50	Matches:	131
Percent Similarity:	79.14%	Conservative:	17
Best Local Similarity:	70.05%	Mismatches:	38
Query Match:	63.26%	Indels:	1
DB:	24	Gaps:	1

US-09-782-953-9 (1-197) x AAD030158 (1-828)

QY 11 SerThleuValAlaCyValAlaPheValGluValPheThrValGluValGlu 30
 DB 161 ACCTCACTTTCTTCTGTCAGCCGTCATGAGCAGTGTTCAGCAGAGAGAGAGAA 220
 QY 31 LysPheGluGlyLeuPheArgThrTyrAspGluCysValThrPheGlnLeuPheYser 50
 DB 221 AGATTGAGACCTCTTCCACCATCTTGATGACACAGTTACTTTACGCTTTAAAC 280
 QY 51 PheArgArgValArgIleAsnPheSerHisProlYserAlaAlaArgAlaArgIleGlu 70
 DB 281 TTTCAGAGAGTCAGATTAATTTTCAGCAACCTGAGAGCGGACAGAGCGCAATAGAA 340
 QY 71 LeuHisGluThrGlnPheArgGlyLeuValLeuLeuLeuTyrPheAlaGlnValGlnThr 90
 DB 341 CTCACGAAACAGACTTCAATGGGAGAGAGCTTAAGCTATATTTTTCACAGGTGACAGATG 400
 QY 91 ProGluThrArgIlePheValPheHisLeuAlaProProGlnProAlaYserGlnPheLeu 110
 DB 401 TCCGCGCAAGTGGGGAGCAAGTCTATCTCTGCGCCGACGCTGCAAGCAGTTCCCTC 460
 QY 111 IleSerProPheSerSerProProValGlyTyrLeuPheProIleSerAlaThrProVal 130
 DB 461 ATCTCCCTCCAGCCTCTCCGCGAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 520
 QY 131 LeuAsnTyrAspLeuLeuTyrAlaValAlaLeuLeuGlyProGlyGlyTyrGluLeu 150
 DB 521 ATAAATATATATTACTCTGCTGCTGTTCCAAATTTGGACAGAGAGAGAGAGAGAGAG 580
 QY 151 HisAlaGlyThrGluSerThrProSerValValHisValCysAspSerAlaPheGlu 170
 DB 581 CACCGGGAACAGATGACACCCAGCGTGTGTTATGTTCTGTAAGTGAAGCACTGAA 640
 QY 171 GluGluGluAspProlYserThrSerProlYserProlYserIleGlnThrArgArgProGly 190
 DB 641 GAGGAAAG 697
 QY 191 LeuProPheSerValSerAsn 197

DB 698 CTCGAGCCGACGCTTGAAAT 718

RESULT 9

ID AAD30155 standard; DNA; 2331 BP.

AC AAD30155;

DT 17-MAY-2002 (first entry)

DE Human MCIP associated DNA #1.

KW Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;
 KW heart failure; cardiomyopathy; heart disease; human; gene; ds.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS /*tag= a
 FT 144..737
 FT /product= "Human MCIP associated protein #1"

XX WO200204491-A2.

XX 17-JAN-2002.

XX 06-JUL-2001; 2001WO-US21662.

XX 07-JUL-2000; 2000US-216601P.

XX 13-FEB-2001; 2001US-0782953.

XX (TEXA) UNIV TEXAS SYSTEM.

XX (WILL/) WILLIAMS S R.

XX (ROTH/) ROTHERMEL B.

XX Williams SR, Rothermel B;

XX WPI; 2002-179698/23.

XX P-PSDB; AAE18914.

XX Claim 72; Page 155-157; 174pp; English.

XX The invention relates to muscle calcineurin interacting proteins (MCIPs)
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical
 CC complex with the catalytic subunit of calcineurin and increased levels
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate
 CC transcription of certain target genes. The invention also relates to
 CC methods for identifying modulators of MCIP binding, expression or
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin
 CC may be used for treating cardiac hypertrophy and heart failure.

CC Antibodies to MCIP can be used in characterizing the MCIP content of
 CC healthy and diseased tissues and subsequently for determining the
 CC presence or absence of cardiomyopathy or as predictor of heart disease.

CC The present sequence is human MCIP associated DNA.

CC Note: This sequence has been described as a promoter in claim 72 of
 CC the specification, however the sequence seems to be a polynucleotide
 CC encoding a MCIP associated protein.

XX Sequence 2331 BP; 630 A; 470 C; 547 G; 684 T; 0 other;

XX Alignment Scores:

Pred. No.:	1.63e-56	Length:	2331
Score:	641.00	Matches:	120
Percent Similarity:	79.68%	Conservative:	29
Best Local Similarity:	64.17%	Mismatches:	30
Query Match:	61.58%	Indels:	8
DB:	24	Gaps:	2

US-09-782-953-9 (1-197) x AAD30155 (1-2331)

QY 11 SerThrLeuValAlaCysValValAspValGluValPheThrAsnGlnGluValLysGlu 30
 Db 174 AGCTCCCTGATTCCTGTCGCAACAGTGTATATCTTCAGCGAAGTGAACACAGGCC 233
 QY 31 LysPheGluGlyLeuPheArgThrTyrAspGlyCysValThrPheGlnLeuPheLysSer 50
 Db 234 AAATTTGAGTCCCTCTTTAGGACGTATGACAAAGGACATCACCTTTTCAGTATTTTAAGAGC 293
 QY 51 PheArgArgValArgIleAsnPheSerHisProLysSerAlaAlaArgAlaArgIleGlu 70
 Db 294 TTCAACAGAGTCAGATTAATCTTCAGCAACCCCTTCCTCCGACGAGATGCCAGGCTCCAG 353
 QY 71 LeuHisGluThrGlnPheArgGlyLysLysLysLysLysLysLysLysLysLysLysLys 90
 Db 354 CTGATAGACTGAGTTCTCGGAAGAAATGAATTATATATTTTGTCTAGACCTTACAC 413
 QY 91 ProGluThrAspGlyAspLysLeuHisLeuAlaProProGlnProAlaLysGlnPheLeu 110
 Db 414 ATAGAAGCTCA-----CACCTGGCTCCGCCAATCCAGACACAAAGCAGTTTCTG 461
 QY 111 IleSerProProSerSerProProValGlyTyrPlyProIleSerAspAlaThrProVal 130
 Db 462 ATCTCCCTCCCGCTCTCCGCGAGTGGATGGAACAAGTGGAGATGCCACCCAGTC 521
 QY 131 LeuAsnTyrAspLeuLeuTyrAlaValAlaLysLeuGlyProGlyGluLysTyrGluLeu 150
 Db 522 ATAAACTATGATCTCTTATATGCCATCTCCAAAGCTGGGGCCAGGGGAAAGTATGAATTG 581
 QY 151 HisAlaGlyThrGluSerThrProSerValValHisValCysAspSerAspMetGlu 170
 Db 582 CACGACGAGTACACACTCCCGAGGTGGTGTCCATGTGTGTAGAGTGTATCAAGAG 641
 QY 171 GluGluGluAspProLysThrSer-----ProLysProLysIleIleGlnThr 186
 Db 642 AAGCAGGAAGAGAGGAATGGAAGAATGAGGAGACTTAAGCCAAAATATCCAGACC 701
 QY 187 ArgArgProGlyLeuProPro 193
 Db 702 AGGAGCGCGAGTACACGCCG 722

RESULT 10
 AAD30159
 ID AAD30159 standard; DNA; 720 BP.
 AC AAD30159;
 XX 17-MAY-2002 (first entry)
 DT Human MCIP associated DNA #5.
 DE Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;
 KW heart failure; cardiomyopathy; heart disease; human; gene; ds.
 KW Homo sapiens.
 OS
 XX
 XX Key Location/Qualifiers
 FT 2..640
 CDS /tag= a
 FT /product= "Human MCIP associated protein #5"
 XX WO200204491-A2.
 XX 17-JAN-2002.
 XX 06-JUL-2001; 2001WO-021662.
 XX 07-JUL-2000; 2000US-216601P.
 XX 13-FEB-2001; 2001US-0782953.
 XX (TEXA) UNIV TEXAS SYSTEM.
 PA (WILL/) WILLIAMS S R.

(ROTH/) ROTHERMEL B.

Williams SR, Rothermel B;

MPI; 2002-179698/23.

P-PSDB; AAE18918.

Screening for modulators of muscle calcineurin interacting protein (MCIP) binding, expression or phosphorylation, useful for treating cardiac hypertrophy or heart failure, comprises mixing MCIP, calcineurin and a test compound -

Disclosure; Page 170-171; 174pp; English.

The invention relates to muscle calcineurin interacting proteins (MCIPs) and nucleic acid molecules encoding such proteins. MCIPs form a physical complex with the catalytic subunit of calcineurin and increased levels of MCIPs correspond to a reduced ability of calcineurin to stimulate transcription of certain target genes. The invention also relates to methods for identifying modulators of MCIP binding, expression or phosphorylation. Inhibitors or promoters of MCIP binding, expression or phosphorylation may be used for treating cardiac hypertrophy and heart failure. Antibodies to MCIP can be used in characterising the MCIP content of healthy and diseased tissues and subsequently for determining the presence or absence of cardiomyopathy or as predictor of heart disease. The present sequence is human MCIP associated DNA described in the invention.

SQ Sequence 720 BP; 196 A; 168 C; 192 G; 164 T; 0 other;

Alignment Scores:

Pred. No.:	2,12e-56	Length:	720
Score:	633.50	Matches:	127
Percent Similarity:	77.01%	Conservative:	17
Best Local Similarity:	67.91%	Mismatches:	32
Query Match:	60.85%	Indels:	11
DB:	24	Gaps:	2

US-09-782-953-9 (1-197) x AAD30159 (1-720)

QY 11 SerThrLeuValAlaCysValValAspValGluValPheThrAsnGlnGluValLysGlu 30
 Db 83 ACCTCACTTTTGTGTCGCGTCCATGAAGCAGTGTGTAGGCACGAGCAGAAGGAA 142
 QY 31 LysPheGluGlyLeuPheArgThrTyrAspGlyCysValThrPheGlnLeuPheLysSer 50
 Db 143 AGATTTGAGCACTCTTACCATCTATGATGACCAAGTTACTTTTTCAGCTGTTAAAAAGC 202
 QY 51 PheArgArgValArgIleAsnPheSerHisProLysSerAlaAlaArgAlaArgIleGlu 70
 Db 203 TTAGAAGAGTCAGATAAATTTTTCAGCAACCTGAAGCGGCGCAGCAAGCGGCAATAGAA 262
 QY 71 LeuHisGluThrGlnPheArgGlyLysLysLysLysLysLysLysLysLysLysLys 90
 Db 263 CTCACGAAACAGACTTCAATGGCGAGAAAGCTAAAGCTATATTTTGCACAG----- 313
 QY 91 ProGluThrAspGlyAspLysLeuHisLeuAlaProProGlnProAlaLysGlnPheLeu 110
 Db 314 -----TCTATCTCTGCGCGCCCGCCCTGTCAAGCAGTCTCTC 352
 QY 111 IleSerProProSerSerProProValGlyTyrPlyProIleSerAspAlaThrProVal 130
 Db 353 ATCTCCCTCCAGCCTCTCCCGAGTGGGTGGAAGCAGACGAGAGATGCCATGCTGTT 412
 QY 131 LeuAsnTyrAspLeuLeuTyrAlaValAlaLysLysLysLysLysLysLysLysLys 150
 Db 413 ATAAATATGATTTACTCTGTGCTGTTTCCAAATTTGGACCGAGAGAGAAATATCAACTT 472
 QY 151 HisAlaGlyThrGluSerThrProSerValValHisValCysAspSerAspMetGlu 170
 Db 473 CACGCGGGAACAGAGTGCACACCCAGCGTGTGTTTCTGATCTGTGAAGAAAGTAACTGAA 532
 QY 171 GluGluGluAspProLysThrSerProLysProLysIleIleGlnThrArgArgProGly 190

PI Tai J;

XX WPI; 2002-010925/01.

XX Composition useful for diagnosis of conditions, disorders or diseases
 PT associated with atherosclerosis, comprises several polynucleotides that
 PT are differentially expressed in foam cell development -

XX Claim 1; Page 112-113; 315pp; English.

XX The present invention relates to the isolation of human polynucleotide
 CC sequences that are differentially expressed during foam cell
 CC differentiation. The polynucleotide sequences of the invention or a
 CC composition comprising these polynucleotides are useful as a high
 CC throughput method for detecting altered expression of one or more
 CC polynucleotides in a sample. The polynucleotides can be used in the
 CC diagnosis of disorders associated with foam cell development such as
 CC atherosclerosis, cerebral stroke, and cardiovascular disorders such as
 CC coronary artery disease. The polynucleotide sequences can also be used
 CC as PCR primers and probes. The polynucleotide sequences of the invention are also
 CC useful in gene therapy. AAS94746-AAS95021 represent the human
 CC polynucleotide sequences of the invention which are differentially
 CC expressed during foam cell differentiation.

XX Sequence 2348 BP; 638 A; 473 C; 550 G; 687 T; 0 other;

Alignment Scores:

Pred. No.: 3.66e-55 Length: 2348
 Score: 628.00 Matches: 120
 Percent Similarity: 79.26% Conservative: 29
 Best Local Similarity: 63.83% Mismatches: 30
 Query Match: 60.33% Indels: 9
 DB: 24 Gaps: 2

US-09-782-953-9 (1-197) x AAS94805 (1-2348)

Qy 11 SerThrLeuValAlaCysValValAspValGluValPheThrAsnGlnGluValLysGlu 30
 Db 181 AGCTCCCTGATGCTGTGGCAAGCAGTGATATCTTCAGCGAAAGTGAACACCGGGCC 240
 Qy 31 LysPheGluGlyLeuPheArgThrTyrAspGluCysValThrPheGlnLeuPheLysSer 50
 Db 241 AAATTTCAGTCCCTCTTTAGACGTATGACAGGACATCACCTTTTCAGTATTTTAAAGC 300
 Qy 51 PheArgValAlaGlnLeuAsnPheSerHisProLysSerAlaAlaArgAlaGlnGlu 70
 Db 301 TTCAACGAGTCAAGATTAACCTTCAGCAACCCCTTCTCCGACGACATGCCAGGCTCCAG 360
 Qy 71 LeuHisGluThrGlnPheArgGlyLysLeuLysLeuTyrPheAlaGlnValGlnThr 90
 Db 361 CTGATTAAGACTGAGTTCTTGGGNAAGGAATGAAGTATATTGCTCAGACCTTACAC 420
 Qy 91 ProGluThrAspGlyAspLysLeuHisLeu-AlaProProGlnProAlaLysGlnPheLe 110
 Db 421 ATAGGAAGCTCA-----CACCTGGGCTCCGCCAAATCCAGACAGGAGTTTCT 468
 Qy 110 uLeSerProProSerSerProProValGlyTrpLysProIleSerAspAlaThrProVa 130
 Db 469 GATCTCCCTCCCGCCCTCTCCCGAGTGGATGGAACAACAGTGAAGATGCCAGCCCACT 528
 Qy 130 lLeuAsnTyrAspLeuLeuTyrAlaValAlaLysLeuGlyProGlyGlyLysTyrGluLe 150
 Db 529 CATAACTATGATCTCTTATATGCCATCTCCAAGCTGGGGCCAGGGGAAAGATATGAAT 588
 Qy 150 uHisAlaGlyThrGluSerThrProSerValValHisValCysAspSerAspMetGl 170
 Db 589 GCACGCGGACTGACACCACTCCACGCGTGGTGGTCCATGTATGTCAGAGTATCAAGA 648
 Qy 170 uGluGluGluAspProLysThrSer-----ProLysProLysIlelleGlnTh 186
 Db 649 GAAGGAGGAAGAGAGAGGAATGAAGAAATGAGGAGACCTTACCAACAAATATTCAGAC 708
 Qy 186 rArgArgProGlyLeuProPro 193

Db 709 CAGGAGCGCGAGTACACGCCG 730

RESULT 13

AAD30156
 ID AAD30156 standard; DNA; 2212 BP.

XX AAD30156;

XX 17-MAY-2002 (first entry)

XX Human MCIP associated DNA #2.

XX Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;
 KW heart failure; cardiomyopathy; heart disease; human; gene; ds.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 25..618

XX /*tag= a

XX /product= "Human MCIP associated protein #2"

XX WO200204491-A2.

XX 17-JAN-2002.

XX 06-JUL-2001; 2001WO-US21662.

XX 07-JUL-2000; 2000US-216601P.

XX 13-FEB-2001; 2001US-0782953.

XX (TEXA) UNIV TEXAS SYSTEM.

XX (WILL/) WILLIAMS S R.

XX (ROTH/) ROTHERMEL B.

XX Williams SR, Rothermel B;

XX WPI; 2002-179698/23.

XX P-PSDB; AAE18915.

XX Screening for modulators of muscle calcineurin interacting protein
 PT (MCIP) binding, expression or phosphorylation, useful for treating
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,
 PT calcineurin and a test compound -

XX Example 1; Page 159-161; 174pp; English.

XX The invention relates to muscle calcineurin interacting proteins (MCIPs)
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical
 CC complex with the catalytic subunit of calcineurin and increased levels
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate
 CC transcription of certain target genes. The invention also relates to
 CC methods for identifying modulators of MCIP binding, expression or
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin
 CC may be used for treating cardiac hypertrophy and heart failure.

CC Antibodies to MCIP can be used in characterising the MCIP content of
 CC healthy and diseased tissues and subsequently for determining the
 CC presence or absence of cardiomyopathy or as predictor of heart disease.

CC The present sequence is human MCIP associated DNA.

CC Note: This sequence has been described as murine MCIP splice variant in
 CC the specification, however the sequence seems to be a polynucleotide
 CC encoding a MCIP associated protein.

XX Sequence 2212 BP; 588 A; 455 C; 529 G; 640 T; 0 other;

Alignment Scores:

Pred. No.: 4.85e-55 Length: 2212
 Score: 626.50 Matches: 121
 Percent Similarity: 74.63% Conservative: 29
 Best Local Similarity: 60.20% Mismatches: 34
 Query Match: 60.18% Indels: 17
 DB: 24 Gaps: 3

US-09-782-953-9 (1-197) x AAD30156 (1-2212)

QY 6 MetAapCYAapVal-----SerThrLeuValAlaCys 16
 DB 13 ATTGACTGGAGATGAGAGGTGAGCTTGACGAGACCTGCCAGCCACCCCTGCTGT 72
 QY 17 ValValAapValGluValPheThrAengIngluValIleGluLeuPheGluLeuPhe 36
 DB 73 CACCTGGACCCCGCGCTGCTTCTGAGACGGCCGTGCCGCGCAATTGAGTCCCTTT 132
 QY 37 ArgThrTyAapGluCysValThrPheGluLeuPheLeuSerPheArgArgValArgIle 56
 DB 133 AGGAGCTATGACAGAGACATCACTTTCAGTATTTAAAGCTTCAACAGCTCAGAAAT 192
 QY 57 AapPheSerHisProLySerAlaAlaArgAlaArgIleGluLeuHisGluThrGlnPhe 76
 DB 193 AACTTCAGCAACCCCTTCTCCGAGCAGATGCCAGGCTCAGCTGATGAAGACTGAGTT 252
 QY 77 ArgGlyLeuValLeuLeuLeuTyPheAlaGlnValGlnThrProGluThrAapGlyAap 96
 DB 253 CTGGGAAAGAAATGAATGATATTTTCTCAGACCTTACATAGAGCTCA----- 306
 QY 97 LysLeuHisLeuAlaProProGlnProAlaLysGlnPheLeuIleSerProSerSer 116
 DB 307 -----CACCTGGCTCCGCCAAATCCAGACAGAGCTTCTGATCTCCCTCCGCTCT 360
 QY 117 ProProValGlyTyPheProIleSerAapAlaThrProValLeuAapTyAapLeuLeu 136
 DB 361 CCGCCGAGGAGAGAAACAAAGTGAAGATGCAACCCAGCTCAATCAATGATCTCTTA 420
 QY 137 TyValValAlaLysLeuGlyProGlyGluLysTyGluLeuHisAlaGlyThrGluSer 156
 DB 421 TATGCATCTCCCAAGCTGGGCGCAGGAAAGTATGCAATGCAACCAAGAGCTACACC 480
 QY 157 ThrProSerValValHisValCysAapSerAapMetGluGluGluAapProLys 176
 DB 481 ACTCCAGAGCTGTGTGCTCATGTATGTAGAGATGATCAAGAGAGAGAGAAAGAGAA 540
 QY 177 ThrSer-----ProLysProLysIleIleGlnThrArgArgProGlyLeuPro 192
 DB 541 ATGGAAAGAAATGAGAGACCTAAGCCAAATATATCCAGACAGAGCGCGAGTACAGC 600

QY 193 Pro 193
 DB 601 CCG 603

RESULT 14
 AAD30152
 ID AAD30152 standard; DNA; 597 BP.
 AC AAD30152;
 AC AAD30152;
 DT 17-MAY-2002 (first entry)
 XX
 DE Human MCIP associated DNA #2.
 DE
 KM Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;
 KM heart failure; cardiomyopathy; heart disease; human; gene; de.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..597
 FT /tag= a
 FT /product= "Human MCIP associated protein #2"
 FT
 PN MO200204491-A2.
 XX
 PD 17-JAN-2002.
 PF 06-JUL-2001; 2001MO-US21662.
 XX

PR 07-JUL-2000; 2000US-216601P.
 PR 13-FEB-2001; 2001US-0782953.
 XX
 PA (TEXA) UNIV TEXAS SYSTEM.
 PA (WILL/) WILLIAMS S R.
 PA (ROTH/) ROTHERMEL B.
 XX
 PI Williams SR, Rothermel B;
 XX
 DR WPI; 2002-179698/23.
 DR P-PSDB; AAE18911.
 XX
 PT Screening for modulators of muscle calcineurin interacting protein
 PT (MCIP) binding, expression or phosphorylation, useful for treating
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,
 PT calcineurin and a test compound -
 XX
 PS Claim 95; Page 147-148; 174pp; English.
 XX
 CC The invention relates to muscle calcineurin interacting proteins (MCIPs)
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical
 CC complex with the catalytic subunit of calcineurin and increased levels
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate
 CC transcription of certain target genes. The invention also relates to
 CC methods for identifying modulators of MCIP binding, expression or
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin
 CC may be used for treating cardiac hypertrophy and heart failure.
 CC Antibodies to MCIP can be used in characterizing the MCIP content of
 CC healthy and diseased tissues and subsequently for determining the
 CC presence or absence of cardiomyopathy or as predictor of heart disease.
 CC The present sequence is human MCIP associated DNA.
 XX
 SQ Sequence 597 BP; 160 A; 170 C; 151 G; 116 T; 0 other;
 XX

Alignment Scores:
 Pred. No.: 1,96e-54 Length: 597
 Score: 613.50 Matches: 119
 Percent Similarity: 75.00% Conservative: 22
 Best Local Similarity: 63.30% Mismatches: 38
 Query Match: 58.93% Indels: 9
 DB: 24 Gaps: 2

US-09-782-953-9 (1-197) x AAD30152 (1-597)

QY 11 SerThrLeuValAlaCysValAapValGluValPheThrAengIngluValIleGlu 30
 DB 31 AGCGCCACCAGCTGCTGCGCACCTGAGCCGCGGTTCGTGAGACGGCTGTGCGGCC 90
 QY 31 LysPheGluGluLeuPheArgThrTyAapGlyCysValThrPheGlnLeuPheSer 50
 DB 91 AAATTGAATCCCTCTTCAGAACATATGACAAAGACACACCTTCCATATTATTAAGGC 150
 QY 51 PheArgArgValArgIleAapPheSerHisProLySerAlaAlaArgAlaArgIleGlu 70
 DB 151 TTCAAAAGCTGTCCGATTAACCTTCAGAACCCCTTATCTGAGCCGAGTCCAGCTGCGG 210
 QY 71 LeuHisGluThrGlnPheArgGlyLysLysLeuLeuTyPheAlaGlnValGlnThr 90
 DB 211 CTGCACAAAGACCGAGTCTCGTGGGAAAGAAATGAAGTGTATTTGCTCAGACTTACAC 270
 QY 91 ProGluThrAapGlyAapLysLeuHisLeuAlaProProGlnProAlaLysGlnPheLeu 110
 DB 271 ATAGGAAGTTCA-----CACCTGGCTCCGCCCAATCCCGCAAAACAGTTCTC 318
 QY 111 IleSerProProSerSerProProValGlyTyPheProIleSerAapAlaThrProVal 130
 DB 319 ATCTCCCTCCGCGCTCTCTCCGTTGCGTGAAGAAAGTGAAGATGCGACCCCGCTC 378
 QY 131 LeuAapTyAapLeuLeuTyAlaValAlaLysLeuGlyProGlyGluLysTyGluLeu 150
 DB 379 ATAAATTACGATCTTTATATGCAATCTCCAAAGCTGGGCGCAGAGAGAAATGATAACCG 438
 QY 151 HisAlaGlyThrGluSerThrProSerValValHisValCysAapSerAapMetGlu 170

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 11:37:42 ; Search time 24.7083 Seconds
(without alignments)
1651.161 Million cell updates/sec

Title: US-09-782-953-3
Perfect score: 1048
Sequence: 1 MEEVDLQLSATIAHLDP.....RPKPKIIQTRRPEYTIHLS 198

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 21:.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_phc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	911	86.9	198	11 Q91WQ4	Q91WQ4 mus musculus
2	839.5	80.1	171	4 Q96R03	Q96R03 homo sapien
3	627.5	59.9	243	11 Q8VIP4	Q8VIP4 mus musculus
4	608.5	58.1	192	11 Q8VIP5	Q8VIP5 mus musculus
5	155	14.8	29	4 Q9H2A1	Q9H2A1 homo sapien
6	139.5	13.3	315	3 Q9P5S9	Q9P5S9 neurospora
7	114.5	10.9	249	3 Q9P4A1	Q9P4A1 cryptococcu
8	94	9.0	318	10 Q9SL96	Q9SL96 arabidopsis
9	94	9.0	318	10 Q8VZ43	Q8VZ43 arabidopsis
10	92	8.8	328	10 Q9LSD3	Q9LSD3 arabidopsis
11	92	8.8	922	4 Q9NTH6	Q9NTH6 homo sapien
12	92	8.8	1455	4 Q9UPV0	Q9UPV0 homo sapien
13	91.5	8.7	1274	10 Q9ZQK2	Q9ZQK2 arabidopsis
14	90.5	8.6	4025	4 Q9NR13	Q9NR13 homo sapien
15	88.5	8.4	680	17 Q8ZY88	Q8ZY88 pyrobaculum
16	87.5	8.3	1477	5 Q9VB52	Q9VB52 drosophila

17	87	8.3	267	11 Q95JH6	Q95JH6 mus musculus
18	86	8.2	781	4 Q9UIY5	Q9UIY5 homo sapien
19	84.5	8.1	719	11 Q61281	Q61281 mus musculus
20	84	8.0	751	5 Q8SX18	Q8SX18 drosophila
21	84	8.0	817	5 Q9VHB4	Q9VHB4 drosophila
22	83	7.9	1173	11 Q63624	Q63624 rattus norv
23	82.5	7.9	241	5 Q9N2W9	Q9N2W9 caenorhabdi
24	82.5	7.9	560	6 Q46411	Q46411 bos taurus
25	82.5	7.9	560	11 Q9D8G6	Q9D8G6 mus musculus
26	82.5	7.9	611	5 Q8WTE9	Q8WTE9 drosophila
27	82	7.8	697	10 Q9CA19	Q9CA19 arabidopsis
28	82	7.8	1128	11 Q88442	Q88442 mus musculus
29	82	7.8	1477	5 Q76931	Q76931 drosophila
30	81.5	7.8	528	17 Q29668	Q29668 archaeoglob
31	81.5	7.8	590	11 Q9CW63	Q9CW63 mus musculus
32	81.5	7.8	952	11 Q8R3C6	Q8R3C6 mus musculus
33	81	7.7	645	11 Q8VDM7	Q8VDM7 mus musculus
34	81	7.7	708	11 Q91YS3	Q91YS3 mus musculus
35	81	7.7	1099	4 Q9Y6N4	Q9Y6N4 homo sapien
36	81	7.7	1312	4 Q9NR59	Q9NR59 homo sapien
37	81	7.7	1343	4 Q9H7N4	Q9H7N4 homo sapien
38	80	7.6	681	4 Q9UFI8	Q9UFI8 homo sapien
39	80	7.6	727	4 Q9BQS9	Q9BQS9 homo sapien
40	80	7.6	817	5 Q9VXK5	Q9VXK5 drosophila
41	79.5	7.6	797	16 Q67651	Q67651 aquifex aeo
42	79.5	7.6	1020	10 Q9LP46	Q9LP46 arabidopsis
43	79.5	7.6	1136	5 Q9XX01	Q9XX01 caenorhabdi
44	79.5	7.6	1146	16 Q910B0	Q910B0 pseudomonas
45	79.5	7.6	1557	16 Q86560	Q86560 streptomyce

ALIGNMENTS

RESULT 1

Q91WQ4 PRELIMINARY; PRT; 198 AA.
ID AC Q91WQ4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Down syndrome critical region homolog 1 (human).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RA Strausberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL, BC013551; AAH13551.1; --
SQ SEQUENCE 198 AA; 22851 MW; F017C68F18ACC187 CRC64;

Query Match 86.9%; Score 911; DB 11; Length 198;
Best Local Similarity 93.1%; Pred. No. 3 3e-82;
Matches 175; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

Qy	11	SATIACHLDRPVFVGLCRAKFESLRTYDKDTTFQYFKSKRVRINFSNPLSAADARLR	70
Db	11	SSLIACVANDVFESETRAKFESLRTYDKDTTFQYFKSKRVRINFSNPLSAADARLR	70
Qy	71	LHKTEFLGKEMKLYFAQTUHGSSHLAPPNDKQFLISPPASPPVGVGWKQVEDATPVIND	130
Db	71	LHKTEFLGKEMKLYFAQTUHGSSHLAPPNDKQFLISPPASPPVGVGWKQVEDATPVIND	130
Qy	131	LLYAIKSLGPGKGYELHAATDTPSVVHVCSDDQENEEEMERMKRPKPKIIQTRRP	190
Db	131	LLYAIKSLGPGKGYELHAATDTPSVVHVCSDDQENEEEMERMKRPKPKIIQTRRP	190
Qy	191	EYTPHLS 198	
Db	191	EYTPHLS 198	

RESULT 2
Q96R03 PRELIMINARY; PRT; 171 AA.
AC Q96R03;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Down syndrome critical region protein 1.
GN DSCR1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Hua F., Wu J., Zhou Y., Zhang B., Peng X., Qiang B., Yuan J.,
Qiang B.;
RL Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF400429; AAK92478.1; -
SQ SEQUENCE 171 AA; 19823 MW; 133001AEADFOBD9 CRC64;

Query Match 80.1%; Score 839.5; DB 4; Length 171;
Best Local Similarity 95.3%; Pred. No. 3.2e-75;
Matches 161; Conservative 3; Mismatches 4; Indels 1; Gaps 1;

QY 30 AKFESLFRYDKDTTFQYKSFPRVIRINSPNSLAADARLRHKTPELGKEMKLYFAQTL 89
DB 4 AKFESLFRYDKDTTFQYKSFPRVIRINSPNSADARLRHKTPELGKEMKLYFAQTL 63
QY 90 HIGSGHLAPNPDQFLISPPASPPVGMKQVEDATPVINYLVAISKLGPEKYLHAA 149
DB 64 HIGSGHLAPNPDQFLISPPASPPVGMKQVEDATPVINYLVAISKLGPEKYLHAA 123
QY 150 TDPTPSVVHVHVCESDQENEESEEMERMKRPKXIIQTRPEYTPHLS 198
DB 124 TDPTPSVVHVHVCESDQENEESEEMERMKRPKXIIQTRPEYTPHLS 171

RESULT 3
Q9VIP4 PRELIMINARY; PRT; 243 AA.
AC Q9VIP4;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Calcineurin inhibitory protein ZAKI-4 beta.
GN ZAKI-4 BETA.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Kanou Y., Miyazaki T., Seo H., Murata Y.;
RT "calcineurin inhibitory protein ZAKI-4."
RL Submitted (MAY-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AB061525; BAB71956.1; -
DR InterPro; IPR000504; RMA_rec_moc.
DR Pfam; PF00076; trm_1.
SQ SEQUENCE 243 AA; 27332 MW; 2CDBQIA810291851 CRC64;

Query Match 59.9%; Score 627.5; DB 11; Length 243;
Best Local Similarity 62.4%; Pred. No. 4.8e-54;
Matches 121; Conservative 24; Mismatches 40; Indels 9; Gaps 2;

QY 5 DLQDLPASATLACHLDPFVGLCRAKFESLFRYDKDTTFQYKSFPRVIRINSPNSLA 64
DB 51 DFLDLPNSLFCVHVQSFVEESKEFGLFRYDVCVTFQYKSFPRVIRINSHPSKA 110
QY 65 ADARLRHKTPELGKEMKLYFAQTLHIGSS---HLAPNPDQFLISPPASPPVGMKQV 120

DB 111 ARARIELHETQFGKKLKYFAQVOTPETDGDKHLAPPOPAKQFLISPPSSPPVGMKPI 170
QY 121 EDATPVINYLVAISKLGPEKYLHAAATDPTPSVVHVHVCESDQENEESEEMERMKRP 180
DB 171 SDATPVINYLVAISKLGPEKYLHAAATDPTPSVVHVHVCESDQENEESEEMERMKRP 225
QY 181 KPXIOTRRPEYTP 194
DB 226 KPXIOTRRPEYTP 239

RESULT 4
Q9VIP5 PRELIMINARY; PRT; 192 AA.
AC Q9VIP5;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Calcineurin inhibitory protein ZAKI-4.
GN ZAKI-4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Kanou Y., Miyazaki T., Seo H., Murata Y.;
RT "calcineurin inhibitory protein ZAKI-4."
RL Submitted (MAY-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AB061524; BAB71955.1; -
SQ SEQUENCE 192 AA; 21540 MW; B2DB9819F4836A8 CRC64;

Query Match 58.1%; Score 608.5; DB 11; Length 192;
Best Local Similarity 62.8%; Pred. No. 2.6e-52;
Matches 118; Conservative 22; Mismatches 39; Indels 9; Gaps 2;

QY 11 SATLACHLDPFVGLCRAKFESLFRYDKDTTFQYKSFPRVIRINSPNSLAADARLR 70
DB 6 STLACVADVDFVFTNQEVKEFGLFRYDVCVTFQYKSFPRVIRINSHPSKARARIR 65
QY 71 LKHTPELGKEMKLYFAQTLHIGSS---HLAPNPDQFLISPPASPPVGMKQVEDATPV 126
DB 66 LKHTPELGKEMKLYFAQVOTPETDGDKHLAPPOPAKQFLISPPSSPPVGMKPISDATPV 125
QY 127 INYLVAISKLGPEKYLHAAATDPTPSVVHVHVCESDQENEESEEMERMKRPKXIIQ 186
DB 126 LKYLVLVAVKLGPEKYLHAAATDPTPSVVHVHVCESDQENEESEEMERMKRPKXIIQ 180
QY 187 TRPEYTP 194
DB 181 TRPEYTP 188

RESULT 5
Q9H2A1 PRELIMINARY; PRT; 29 AA.
AC Q9H2A1;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE DSCR1 (Fragment).
GN DSCR1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Ernak G., Morgan T., Davies K.J.A.;
RT "Differential expression of adap78 (DSCR1) gene isoforms 1 and 2 in
human tissues, with particular reference to neurons."
RL Submitted (SEP-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF303449; AAG40774.1; -
FT NON_TER 29 29

SQ SEQUENCE 29 AA; 3244 MW; D8CB7D1CD1521082 CRC64;

Query Match 14.8%; Score 155; DB 4; Length 29;
Best Local Similarity 100.0%; Pred. No. 1.6e-08;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEEVDLQDLPSTAIACHLDPRVFDGLCR 29
DB 1 MEEVDLQDLPSTAIACHLDPRVFDGLCR 29

RESULT 6

ID Q9P5S9 PRELIMINARY; PRT; 315 AA.
AC Q9P5S9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Nebula related protein.
GN B5022.230.

OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariales; Sordariales; Neurospora.
OC NCBI_TaxID=5141;
RN [1]

RP SEQUENCE FROM N.A.
RA Schulte U., Aign V., Hoheisel J., Brandt P., Partmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
[2]

RP SEQUENCE FROM N.A.
RA German Neurospora genome project;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; ALJ55932; CAB91442.1; -;
SQ SEQUENCE 315 AA; 34194 MW; 76792987653C3AB7 CRC64;

Query Match 13.3%; Score 139.5; DB 3; Length 315;
Best Local Similarity 40.2%; Pred. No. 1.2e-05;
Matches 39; Conservative 13; Mismatches 38; Indels 7; Gaps 3;

QY 49 KSPKRVINFSNPLSAADARLRLHKTFLGKEMKLYPAQTLHGSS----HLAPPNPDQK 104
DB 101 KSFRIITVFDEQAALAVRSMVDEAILGRCRVYFGPTPIDVSAADKHLALPDAGKL 160

QY 105 FLISPPASPPGVKQ-VEDA--TPVINYDLYAISKL 138
DB 161 FFI5PPSPPHDMEQMEDAPNTWHAEDLAELAKL 197

RESULT 7

ID Q9P4A1 PRELIMINARY; PRT; 249 AA.
AC Q9P4A1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Calcineurin-binding protein.
GN CBP1

OS Cryptococcus neoformans var. neoformans.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
OC NCBI_TaxID=40410;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=H99;
RX MEDLINE=20359261; PubMed=10899116;

RA Gorlach J., Fox D.S., Cutler N.S., Cox G.M., Perfect J.R., Heitman J.;
RT Identification and characterization of a highly conserved calcineurin
binding protein, CBP1/calciressin, in Cryptococcus neoformans.;
RL EMBO J 19:3618-3629(2000).
DR EMBL; AF230799; AAF63734.1; -;

SQ SEQUENCE 249 AA; 27293 MW; 7B757183FA5EBF09 CRC64;

Query Match 10.9%; Score 114.5; DB 3; Length 249;
Best Local Similarity 27.9%; Pred. No. 0.0027;
Matches 46; Conservative 23; Mismatches 61; Indels 35; Gaps 8;

QY 10 PSATIACHL-DRPVF---VDGLCRKPFESLRTYDKDTTFQYFKSFVRVNFNSPLSAA 65
DB 18 PTNTLALLPHPTLFAPPVLDLLRAHYEH---FGRIVHWAPVRGFGRAIVVFSEEEAE 73

QY 66 DAR-----LRL-----HKTEFLGKEMKLYPAQTLHI-----GSSHLAPPNPK 103
DB 74 NAKRQGDWLKLDVPVGGEEKVQNEGKDIELVRLHLHLPPTLTNPDPATHTLAPPPLPH 133

QY 104 QFLISPPASPPGVKQVEDATP---VINYDLYAIS--KLGPCK 143
DB 134 NFLISPPGSPGCEWPAEAPNRILPEDLQRALETLELNSGSK 178

RESULT 8

ID Q9SL96 PRELIMINARY; PRT; 318 AA.
AC Q9SL96;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE At2G25670 protein.
GN AT2G25670.

OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. COLUMBIA;

RX MEDLINE=20083487; PubMed=10617197;

RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA Buel C.R., Schum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Chen M., VanAken S.E., Umavam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carrara A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Preuss D., Niernan W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
thaliana.";
RL Nature 402:761-768(1999).
RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. COLUMBIA;

RA Lin X.;

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AC006053; AAD31367.1; -;

SQ SEQUENCE 318 AA; 35128 MW; 2F2E2CD269FFD004 CRC64;

Query Match 9.0%; Score 94; DB 10; Length 318;

Best Local Similarity 25.2%; Pred. No. 0.39;

Matches 38; Conservative 19; Mismatches 64; Indels 30; Gaps 5;

QY 54 VRINFSNPLSAADARLRLHKTFLGKEMKLYPAQTLHGSSHLAPP--NPDQFLISPPA 111
DB 13 ITIQSTNIFALDTRKKKKSKXAGSK-----GSSKSRPEKEPEQVYWPATP 62

QY 112 SPVGVKQVEDATPVINYDLYAISKLGPCKYELHAATDTPPSVVHVHVCESDQE----- 166
DB 63 LKVKSWADIDDDDDDDVYATTAPQSGWSTLPSH-----TDSKDVHVESESEDILD 117

QY 167 -----NEEEHEEMRMKRPKPIQTRPE 191

DB 118 EGDDVVEEQEETEVOVHPPEV--KKAPE 146

RESULT 9

ID Q8VZ43 PRELIMINARY; PRT; 318 AA.

AC Q8V243;
 DT 01-MAR-2002 (TReMBLrel. 20, Created)
 DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
 DE Hypothetical 35.2 kDa protein.
 GN ATG25670.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsi.
 OC NCBI_TaxID=3702;
 OX [1]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Banh J., Chang C.H., Chang E., Dale J.M., Goldsmith A.D.,
 Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M., Wu H.C.,
 Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Chou R.,
 Hayashizaki Y., Ishida J., Jones T., Kamuya A., Karlin-Neumann G.,
 Kawai J., Kim C., Lam B., Lin J., Meyers M.C., Miranda M.,
 Narusaka M., Nguyen M., Palm C.J., Sakurai T., Saitou M., Seki M.,
 Shinn P., Southwick A., Shinzaki K., Davis R.W., Ecker J.R.,
 Theologis A.,
 RT "Full Length cDNA of gene At2g25670 (GI:15225169)."
 RL Submitted (Dec-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY065267; AL38743.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 318 AA; 35154 MW; 2F2E2CC278FED004 CRC64;
 Query Match 9 0%; Score 94; DB 10; Length 318;
 Best Local Similarity 25.2%; Pred. No. 0.39; Indels 30; Gaps 5;
 Matches 38; Conservative 19; Mismatches 64;
 QY 54 VRIINFSNPADARLRLKTEFLGKEMKLYFAQTLLIGSSHLAPP--NPDKQFLISPPA 111
 DB 13 IITGSNLFALDTRKKKKSKDKAGSK-----GSSKSRBEKPEPQVVAAPRP 62
 QY 112 SPVGVKQVEDATPVINYDLVAISLKGGEKYLHAATDPPIPVVYVVCESDQF----- 166
 DB 63 LKVKSWADIDDEDDDDVATTAPOSQWSTLSLPSH-----TDSKDVHVEESEEDILD 117
 QY 167 -----NEEEEMERMKRPKPIIOTRRPE 191
 DB 118 EGDVDVEEOEETEVQVHPEPV--KKAPE 146
 RESULT 10
 Q9LSD3 PRELIMINARY; PRT; 328 AA.
 ID Q9LSD3;
 AC Q9LSD3;
 DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
 DE Genomic DNA, chromosome 3, pi clone: MOJ10.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsi.
 OC NCBI_TaxID=3702;
 OX [1]
 RP SEQUENCE FROM N.A.
 RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamitsu E., Tabata S.;
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA STRAIN=COLUMBIA;
 RX MEDLINE=20277480; PubMed=10819329;
 RA Nakamura Y.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
 features of the regions of 4,504,864 bp covered by sixty P1 and TAC
 clones."
 DT DNA Rep. 7; 131-135 (2000).
 RL EMBL; AB026649; BAB01084.1; -
 DR EMBL; IPR001776; Aetolysin.

DR PROSITE; PS00274; AEROLYSIN; UNKNOWN.1.
 SQ SEQUENCE 328 AA; 37438 MW; A43A5FEIAD0634AA CRC64;
 Query Match 8.8%; Score 92; DB 10; Length 328;
 Best Local Similarity 21.4%; Pred. No. 0.64; Indels 84; Gaps 9;
 Matches 50; Conservative 24; Mismatches 76;
 QY 23 FVDGLCRAKESLERTYKDTTFQYFKSPKRVIN-PS-----NPLSADARLRLK 73
 DB 31 FKGLGCLR-----AQVFSVDQDIRTSFSGYTHDRNP-SKVQDLRFDE 76
 QY 74 TELFGKEMKLYFAQTLLIGSSHLAPPND-----KQFLIS 108
 DB 77 DEFG-----FLAIGTLGTDPETPKFSAMVAEEDATGEIKENAKLIAKLDOFLKE 127
 QY 109 PPASPVMKQVEDATPVINYDL-----LVNISKLGPEKY 144
 DB 128 YPEDTRSKRVKISNECFLQDYDLFRSIELTKSGNKGKKSLTSLFKRQTVGEPT 167
 QY 145 -ELHAATDPPIPVV--VHVCESDQNEEEEMERMKRPKPIIOTRRPEYTP 195
 DB 188 IEKSTRDAIKRVFKLHGASSKTRNDEDDSMGKKKDKKNVOTCRKRVHPV 241
 RESULT 11
 Q9NTH6 PRELIMINARY; PRT; 992 AA.
 ID Q9NTH6;
 AC Q9NTH6;
 DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
 DE Hypothetical 112.6 kDa protein (fragment).
 GN DKP2P43402413.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxID=9606;
 OX [1]
 RP SEQUENCE FROM N.A.
 RA TISSUE=TESTIS;
 RA Koehler K., Beyer A., Mewes H.W., Gassenhuber J., Wiemann S.;
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL137265; CAB70664.1; -
 KW Hypothetical protein.
 FT NON-TER
 SQ SEQUENCE 992 AA; 112628 MW; 039DF5B1E97E02F CRC64;
 Query Match 8.8%; Score 92; DB 4; Length 992;
 Best Local Similarity 22.7%; Pred. No. 2.8;
 Matches 39; Conservative 18; Mismatches 39; Indels 76; Gaps 7;
 QY 96 LAPNPDKQFLISPPAS-----PPVG-----WKQVEDATPVINYDLVAIS----- 136
 DB 245 ISPLPHERAQSPPRSILATEEPFGQEPGEWEKAE-----LGEQSAASLSQLSLQR 300
 QY 137 -----KLGRG-----EKYELHAATDPPIPVV----- 156
 DB 301 EQAPSPACKEGKQSHQAEELGPGQEADEBEKVAVSPPIPVVSTPEVSTPEVAPPE 360
 QY 157 -----VHVCESDQNEEEEMERMKRPKPIIOTRRPEYTP 196
 DB 361 LSEALKAMEAVAVQVLEDDQKHLLSKQ-EKMQQLRKLQEEBEELRLH 411
 RESULT 12
 Q9UPV0 PRELIMINARY; PRT; 1455 AA.
 ID Q9UPV0;
 AC Q9UPV0;
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
 DE KIAA1052 protein.
 GN KIAA1052.

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BRAIN;
 EX MEDLINE=99397452; PubMed=10470851;
 RA Kikuno A., Nagase T., Ishikawa K., Hiroseawa M., Miyajima N.,
 RA Tanaka A., Kotani H., Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XIV.
 RT The complete sequences of 100 new cDNA clones from brain which code
 RT for large proteins in vitro.";
 RL DNA Res. 6:197-205(1999).
 DR EMBL; AB028975; BAA83004.1; -
 DR InterPro; IPR001202; WW_Rep5_WWP.
 DR Pfam; PF00397; WW; 1.
 DR SMART; SM00456; WW; 1.
 DR PROSITE; PS50020; WW DOMAIN 2; 1.
 SQ SEQUENCE 1455 AA; 163543 MW; 7F48093100C34819 CRC64;
 Query Match 8.8%; Score 92; DB 4; Length 1455;
 Best Local Similarity 22.7%; Pred. No. 4.5;
 Matches 39; Conservative 18; Mismatches 39; Indels 76; Gaps 7;
 QY 96 LAPNPDKQFLISPPAS-----PPVG-----WKVEDATPVINYLLVAIS----- 136
 Db 474 LSPPLPHEAQSPPRSRLATEEPPQPEQPEWKEAEE-----LGEDSAASLSLSLQSR 529
 QY 137 -----KLGPG---EKYELHAATDTPSV----- 156
 Db 530 EQAPSPPAECCKGEQHSQAEELGPGGEAEDPEEKVAVSTPPVSPSEVSTEPVAPPEQ 589
 QY 157 -----VVHVCSDQENEEEMERKRPKPIQTRRPVYTPH 196
 Db 590 LSEAAKAMEEAVQVLEQDQRLHLESQ-EKMQQLREKLQCEEEELRLH 640
 RESULT 13
 Q9ZQK2 ID Q9ZQK2 PRELIMINARY; PRT; 1274 AA.
 AC Q9ZQK2;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Putative retroelement pol polyprotein.
 GN Ar2G07400.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC euroids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RX MEDLINE=20083487; PubMed=10617197;
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
 RA Buell C.R., Ketchum K.A., Lee J.J., Renning C.M., Koo H., Moffat K.S.,
 RA Cronin L.A., Shen M., VanAken S.E., Unayam L., Tallon L.J., Gill J.E.,
 RA Adams M.D., Carreira A.J., Creasy T.H., Goodman H.M., Smerville C.R.,
 RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
 RA Salzberg S.L., Fraser C.M., Venter J.C.;
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
 RT thaliana.";
 RL Nature 402:761-768(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RX MEDLINE=20083487; PubMed=10617197;
 RA Lin X.;
 RP Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AC006217; AACD1532.1; -
 DR InterPro; IPR004312; ATHILA.

DR InterPro; IPR003653; SUMO_protease.
 DR Pfam; PF03078; ATHILA_1
 DR Pfam; PF02902; Peptidase_C48; 1.
 KW Polyprotein.
 SQ SEQUENCE 1274 AA; 145232 MW; A2342B5FDCE45631 CRC64;
 Query Match 23.7%; Score 91.5; DB 10; Length 1274;
 Best Local Similarity 23.7%; Pred. No. 4.3;
 Matches 51; Conservative 30; Mismatches 83; Indels 51; Gaps 9;
 QY 24 VDGLCRKAFSLRPTVDKDTTFQYFKSKRVRINFNPLSAADAR-LRLHKTFPLGKE-- 80
 Db 673 IESRVEAKFEGFRGSIENDV-----KQIKQLKAIADSKSSVIRDMYLAKTOPQTQDN 727
 QY 81 MKLYFAQTLHI-----GSSHLAPNPD-----KQFLISPPA-----SPVGVWK 118
 Db 728 PKVQTQTPDVPKPTNNQFATSPPPSKQADVGVKKKTKTINELIQPKPRGRGKPSQPK 787
 QY 119 QVEDATPVINYLLVAISKLGPGKVELHA-----ATDTPSVVVHVCSDQENEEEEE 173
 Db 788 KVSPTWPKI-----TIKKLPQSEAKAEEDSDVVVDTKVSEYNEMLPSEDEDEEE 841
 QY 174 MERMKRPKPIQTRR-----PEYTPHLS 198
 Db 842 TERLKSVMKIRLTKVTKLAPDGRSLINPEAMPHTHS 876
 RESULT 14
 Q9NR13 ID Q9NR13 PRELIMINARY; PRT; 4025 AA.
 AC Q9NR13;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE ALR-like protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=CERVICAL CARCINOMA;
 RX Chow V.T.K., Tan Y.C.;
 RT "ALR-like protein, a novel human cDNA whose product is homologous to
 RT the ALR protein.";
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: CONTAINS 1 SET DOMAIN.
 DR EMBL; AF264750; AAF74766.2; -
 DR InterPro; IPR000194; ATPase_a/bcentre.
 DR InterPro; IPR000345; CytC heme bind.
 DR InterPro; IPR002219; DAG P8-bind.
 DR InterPro; IPR003889; FYIich C.
 DR InterPro; IPR003888; FYIich N.
 DR InterPro; IPR000910; HMG 12 box.
 DR InterPro; IPR003616; PostSET.
 DR InterPro; IPR001214; SET.
 DR InterPro; IPR001965; Znf PHD.
 DR Pfam; IPR001841; Znf-ring.
 DR Pfam; PF00505; HMG_box; 1.
 DR Pfam; PF00628; PHD; 3.
 DR Pfam; PF00856; SET; 1.
 DR SMART; SM00109; C1; 1.
 DR SMART; SM00542; FYRC; 1.
 DR SMART; SM00541; FYRN; 1.
 DR SMART; SM00398; HMG; 1.
 DR SMART; SM00249; PHD; 4.
 DR SMART; SM00508; PostSET; 1.
 DR SMART; SM00184; RING; 3.
 DR SMART; SM00317; SET; 1.
 DR PROSITE; PS00152; ATPASE ALPHA BETA; UNKNOWN 1.
 DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN 1.
 DR PROSITE; PS50280; SET; 1.
 SQ SEQUENCE 4025 AA; 443532 MW; 41624149C28E4BDE CRC64;

Query Match 8.6%; Score 90.5; DB 4; Length 4025;

Best Local Similarity 27.9%; Pred. No. 24;
Matches 31; Conservative 15; Mismatches 46; Indels 19; Gaps 4;

QY . 87 QTLHGSHLAPNDKOP--LISPPAPVG--NKQVEDATPVINYDLIAISKLGPE 142
Db 941 QVSPGSSNSRPPSPMDPYAKWGTFRPPVGHSPSRNSAPVENCPTLSSVSR----- 995
QY 143 KYELHAATDPTFSVVVHVCSQDENEEEMERKRPKRIOTRPEYT 193
Db 996 PLQMETTANRPSVPRDLCSSTTND-----PYAKPPDTPRPVMT 1036

RESULT 15

Q82Y88 PRELIMINARY; PRT; 680 AA.
ID Q82Y88
AC Q82Y88;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE DNA replication licensing factor (mcm).
GN PAE0901.
OS Pyrobaculum aerophilum.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Pyrobaculum.
OX NCBI_TaxID=13773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
RX PubMed=1192869;
RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
RA Miller J.H.;
RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
aerophilum";
RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
DR EMBL; AE009792; AAL63108.1; -.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001208; MCM_
DR Pfam; PF00493; MCM; 1.
DR ProDom; PD001041; MCM; 1.
DR SMART; SM00382; AAA; 1.
DR SMART; SM00350; MCM; 1.
DR PROSITE; PS50051; MCM_2; 1.
KW Complete proteome.
SQ SEQUENCE 680 AA; 76173 MW; 79409981D729D151 CRC64;

Query Match 8.4%; Score 88.5; DB 17; Length 680;

Best Local Similarity 23.6%; Pred. No. 3.7;
Matches 50; Conservative 30; Mismatches 59; Indels 73; Gaps 12;

QY 40 DKDTQGYKSPKRVINFSNPLSADARLRLHKTFLGKMKL----- 83
Db 80 DEPTAKALKRFFRVR--GSPLSVS---LRKLSEYIGRLIKIEGIVTRQTPPKHFLYK 133
QY 84 -----YFAQTLHGSHLAP-----NPDKQFLISPPASPPVGMKV-----ED 122
Db 134 ALYRCQCGEIEILLQELRHEVPEPAKPCRCGASKSFLLVTELSQYIDWQKVIQGERPD 193
QY 123 ATP-----VINVDLVAISKLGEGEKELHAATDPT-----PSVV-----VHV 160
Db 194 LPPGQLPRSVYVLLDDL--VDTVPKPDIIISLGVVDLTLSELKGRPPIVTSYIGVHV 251
QY 161 CESDQENEEEMERKRPKRIIO--TRPE 191
Db 252 ---DTMKKELVEI--TKEDQKILKISRPD 278

Search completed: December 11, 2002, 11:41:02
Job time : 28.7063 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 11:37:42 ; Search time 30.7184 Seconds
(without alignments)
858.887 Million cell updates/sec

Title: US-09-782-953-3
Perfect score: 1048
Sequence: 1 MEEDVLDLPSATIACHLDP.....RPKPKIIQTRPEYTPPIHLS 198

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	1048	100.0	198	23 AAE18911 Human MCIP associa
2	1002.5	95.7	197	23 AAE18915 Human MCIP associa
3	920.5	87.8	198	23 AAE18912 Mouse MCIP associa
4	872.5	83.3	197	23 AAE18914 Human MCIP associa
5	835.5	79.7	171	20 AA073898 Human Down's Syndr
6	620.5	59.2	234	20 AAB31788 Amino acid sequenc
7	620.5	59.2	255	23 AAMS0760 Human Down syndrom
8	613.5	58.5	197	23 AAE18913 Mouse MCIP associa
9	596.5	56.9	192	23 AAE18916 Human MCIP associa
10	596	56.9	241	23 AAE18917 Human MCIP associa

11	595	56.8	212	23 AAE18918 Human MCIP associa
12	497	47.4	142	21 AAB58452 Lung cancer associ
13	439	41.9	111	21 AAG01768 Human secreted pro
14	379.5	36.2	292	22 ABB71467 Drosophila melanog
15	284	27.1	58	22 ABB29495 Peptide #2146 enco
16	284	27.1	58	22 ABB34672 Peptide #2178 enco
17	284	27.1	58	22 ABB36877 Peptide #4383 enco
18	284	27.1	58	22 ABB20081 Protein #2080 enco
19	284	27.1	58	22 AAM55464 Human brain expres
20	284	27.1	58	22 AAM57623 Human brain expres
21	284	27.1	58	22 AAM67851 Human bone marrow
22	284	27.1	58	22 AAM15671 Peptide #2105 enco
23	284	27.1	58	22 AAM17859 Peptide #4293 enco
24	284	27.1	58	22 AAM28174 Peptide #2211 enco
25	284	27.1	58	22 AAM30365 Peptide #4402 enco
26	284	27.1	58	22 AAM03406 Peptide #2088 enco
27	284	27.1	58	22 AAM05506 Peptide #4188 enco
28	284	27.1	58	23 AAB37397 Human peptide enco
29	284	27.1	58	23 ABB39657 Human peptide enco
30	271.5	25.9	56	22 ABB28178 Human peptide #829
31	271.5	25.9	56	22 ABB33353 Peptide #859 enco
32	271.5	25.9	56	22 ABB18813 Protein #812 enco
33	271.5	25.9	56	22 AAM54139 Human brain expres
34	271.5	25.9	56	22 AAM66532 Human bone marrow
35	271.5	25.9	56	22 AAM14405 Peptide #839 enco
36	271.5	25.9	56	22 AAM26818 Peptide #855 enco
37	271.5	25.9	56	22 AAM02132 Peptide #814 enco
38	271.5	25.9	56	23 ABB36184 Human peptide enco
39	98	9.4	19	23 AAE18921 Arabidopsis thalia
40	94	9.0	318	21 AAG30396 Arabidopsis thalia
41	94	9.0	318	21 AAG44378 Arabidopsis thalia
42	94	9.0	318	21 AAG49385 Arabidopsis thalia
43	94	9.0	318	21 AAG49386 Arabidopsis thalia
44	92	8.8	954	22 AAU14615 Novel bone marrow
45	90.5	8.6	4019	22 AAE13839 Human lung tumour-

ALIGNMENTS

RESULT 1
AAE18911
ID AAE18911 standard; Protein; 198 AA.
AC AAE18911;
XX
DT 17-MAY-2002 (first entry)
XX
DE Human MCIP associated protein #2.
XX
KW Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;
XX heart failure; cardiomyopathy; heart disease; human.
OS Homo sapiens.
XX
PN WO200204491-A2.
XX
PD 17-JAN-2002.
XX
PF 06-JUL-2001; 2001WO-US21662.
XX
PR 07-JUL-2000; 2000US-216601P.
PR 13-FEB-2001; 2001US-0782953.
XX
PA (TEXA) UNIV TEXAS SYSTEM.
PA (WILL/) WILLIAMS S R.
XX (ROTH/) ROTHERMEL B.
PI Williams SR, Rothermel B;
XX
DR WPI; 2002-179698/23.
DR N-PSDB; AAD30152.
XX

PT Screening for modulators of muscle calcineurin interacting protein
 PT (MCIP) binding, expression or phosphorylation, useful for treating
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,
 PT calcineurin and a test compound -

XX Discloure; Page 148-149; 174pp; English.

PS The invention relates to muscle calcineurin interacting proteins (MCIPs)
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical
 CC complex with the catalytic subunit of calcineurin and increased levels
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate
 CC transcription of certain target genes. The invention also relates to
 CC methods for identifying modulators of MCIP binding, expression or
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin
 CC may be used for treating cardiac hypertrophy and heart failure.
 CC Antibodies to MCIP can be used in characterizing the MCIP content of
 CC healthy and diseased tissues and subsequently for determining the
 CC presence or absence of cardiomyopathy or as predictor of heart disease.
 CC The present sequence is human MCIP associated protein.

XX Sequence 198 AA:

Query Match 100.0%; Score 1048; DB 23; Length 198;
 Best Local Similarity 100.0%; Pred. No. 1.6e-108;
 Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEEVDLQDLPSSATLACHDPRVFDGLCRAPESLFRITDQDTFOYFKSKRVIRINFSN 60
 DB 1 MEEVDLQDLPSSATLACHDPRVFDGLCRAPESLFRITDQDTFOYFKSKRVIRINFSN 60
 QY 61 PLSAADARLRHLKTEFLGKEMKLYPAQTLLHIGSSHLAPPNDKQFLISPPASPPVGMKQV 120
 DB 61 PLSAADARLRHLKTEFLGKEMKLYPAQTLLHIGSSHLAPPNDKQFLISPPASPPVGMKQV 120
 QY 121 EDATPVINVDLVAISKLGPEKVELHAATDTPSVVVHVCSDOENEEEMERMRKP 180
 DB 121 EDATPVINVDLVAISKLGPEKVELHAATDTPSVVVHVCSDOENEEEMERMRKP 180
 QY 181 KPRTIOTRRPEYTPHLS 198
 DB 181 KPRTIOTRRPEYTPHLS 198

RESULT 2

AAE18915
 ID AAE18915 standard; Protein; 197 AA.

AC AAE18915;

DT 17-MAY-2002 (first entry)

DE Human MCIP associated protein #2.

KW Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;
 heart failure; cardiomyopathy; heart disease; human.

OS Homo sapiens.

PN W0200204491-A2.

PD 17-JAN-2002.

PF 06-JUL-2001; 2001WO-US21662.

PR 07-JUL-2000; 2000US-216601P.

PR 13-FEB-2001; 2001US-0782953.

XX (TEXTA) UNIV TEXAS SYSTEM.
 XX (WILL) WILLIAMS S R.
 XX (ROTH) ROTHERMEL B.

PI Williams SR, Rothermel B;

XX

DR WPI; 2002-179698/23.
 DR N-PSDB; AAD30156.

XX Screening for modulators of muscle calcineurin interacting protein
 PT (MCIP) binding, expression or phosphorylation, useful for treating
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,
 PT calcineurin and a test compound -

XX Example 1; Page 161-162; 174pp; English.

PS The invention relates to muscle calcineurin interacting proteins (MCIPs)
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical
 CC complex with the catalytic subunit of calcineurin and increased levels
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate
 CC transcription of certain target genes. The invention also relates to
 CC methods for identifying modulators of MCIP binding, expression or
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin
 CC may be used for treating cardiac hypertrophy and heart failure.
 CC Antibodies to MCIP can be used in characterizing the MCIP content of
 CC healthy and diseased tissues and subsequently for determining the
 CC presence or absence of cardiomyopathy or as predictor of heart disease.
 CC The present sequence is human MCIP associated protein.
 CC Note: This sequence has been described as murine MCIP splice variant in
 CC the specification, however the sequence seems to be a MCIP associated
 CC protein.

XX Sequence 197 AA:

Query Match 95.7%; Score 1002.5; DB 23; Length 197;
 Best Local Similarity 96.5%; Pred. No. 1.9e-103;
 Matches 191; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

QY 1 MEEVDLQDLPSSATLACHDPRVFDGLCRAPESLFRITDQDTFOYFKSKRVIRINFSN 60
 DB 1 MEEVDLQDLPSSATLACHDPRVFDGLCRAPESLFRITDQDTFOYFKSKRVIRINFSN 60
 QY 61 PLSAADARLRHLKTEFLGKEMKLYPAQTLLHIGSSHLAPPNDKQFLISPPASPPVGMKQV 120
 DB 61 PLSAADARLRHLKTEFLGKEMKLYPAQTLLHIGSSHLAPPNDKQFLISPPASPPVGMKQV 120
 QY 121 EDATPVINVDLVAISKLGPEKVELHAATDTPSVVVHVCSDOENEEEMERMRKP 180
 DB 121 EDATPVINVDLVAISKLGPEKVELHAATDTPSVVVHVCSDOENEEEMERMRKP 179
 QY 181 KPRTIOTRRPEYTPHLS 198
 DB 180 KPRTIOTRRPEYTPHLS 197

RESULT 3

AAE18912
 ID AAE18912 standard; Protein; 198 AA.

AC AAE18912;

DT 17-MAY-2002 (first entry)

DE Mouse MCIP associated protein #3.

KW Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;
 heart failure; cardiomyopathy; heart disease; mouse.

OS Mus musculus.

PN W0200204491-A2.

PD 17-JAN-2002.

PF 06-JUL-2001; 2001WO-US21662.

PR 07-JUL-2000; 2000US-216601P.

PR 13-FEB-2001; 2001US-0782953.

XX

PA (TEXA) UNIV TEXAS SYSTEM.
 PA (WILL/) WILLIAMS S R.
 PA (ROTH/) ROTHERMEL B.
 XX Williams SR, Rothermel B;
 PI WPI: 2002-179698/23.
 XX N-PSDB; AAD30153.
 DR Screening for modulators of muscle calcineurin interacting protein
 XX (MCIP) binding, expression or phosphorylation, useful for treating
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,
 PT calcineurin and a test compound -
 XX Disclosure; Page 151-152; 174pp; English.
 XX The invention relates to muscle calcineurin interacting proteins (MCIPs)
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical
 CC complex with the catalytic subunit of calcineurin and increased levels
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate
 CC transcription of certain target genes. The invention also relates to
 CC methods for identifying modulators of MCIP binding, expression or
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin
 CC may be used for treating cardiac hypertrophy and heart failure.
 CC Antibodies to MCIP can be used in characterising the MCIP content of
 CC healthy and diseased tissues and subsequently for determining the
 CC presence or absence of cardiomyopathy or as predictor of heart disease.
 CC The present sequence is mouse MCIP associated protein.
 CC Note: This sequence has been described as human MCIP3 in the
 CC specification, however the sequence seems to be a MCIP associated
 CC protein.
 XX Sequence 198 AA;
 SQ
 Query Match 87.8%; Score 920.5; DB 23; Length 198;
 Best Local Similarity 89.9%; Pred. No. 2.8e-94;
 Matches 178; Conservative 5; Mismatches 12; Indels 3; Gaps 1;
 QY 4 VDLQDLP--SATIACHLDPRVFDGLCRKAFESLRTYDKDTTFQYFKSRVRINFN 60
 DB 1 MDRDFSYNFSLLIACVANDDDVSESETRAKFESLRTYDKDTTFQYFKSRVRINFN 60
 QY 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHIGSSHLAPPNDKQFLISPPASPPVGMKQV 120
 DB 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHIGSSHLAPPNDKQFLISPPASPPVGMKQV 120
 QY 121 EDATPVINVDLLVAISKLGPGKGYELHAATDPTPSVVHVHVCSDQNEEEEEEMERMKRP 180
 DB 121 EDATPVINVDLLVAISKLGPGKGYELHAATDPTPSVVHVHVCSDQNEEEEEEMERMKRP 180
 QY 181 KPXIOTRRPEYTPIHLS 198
 DB 181 KPXIOTRRPEYTPIHLS 198
 RESULT 4
 ID AAE18914
 XX AAE18914 standard; Protein; 197 AA.
 AC AAE18914;
 XX
 DT 17-MAY-2002 (first entry)
 XX Human MCIP associated protein #1.
 DE
 XX Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;
 KW heart failure; cardiomyopathy; heart disease; human.
 XX Homo sapiens.
 OS
 XX WO200204491-A2.
 PN
 XX 17-JAN-2002.
 PD

XX 06-JUL-2001; 2001WO-US21662.
 PF
 XX 07-JUL-2000; 2000US-216601P.
 PR 13-FEB-2001; 2001US-0782953.
 XX
 PA (TEXA) UNIV TEXAS SYSTEM.
 PA (WILL/) WILLIAMS S R.
 PA (ROTH/) ROTHERMEL B.
 XX Williams SR, Rothermel B;
 PI WPI: 2002-179698/23.
 XX N-PSDB; AAD30155.
 DR Screening for modulators of muscle calcineurin interacting protein
 XX (MCIP) binding, expression or phosphorylation, useful for treating
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,
 PT calcineurin and a test compound -
 XX Example 1; Page 157-158; 174pp; English.
 XX The invention relates to muscle calcineurin interacting proteins (MCIPs)
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical
 CC complex with the catalytic subunit of calcineurin and increased levels
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate
 CC transcription of certain target genes. The invention also relates to
 CC methods for identifying modulators of MCIP binding, expression or
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin
 CC may be used for treating cardiac hypertrophy and heart failure.
 CC Antibodies to MCIP can be used in characterising the MCIP content of
 CC healthy and diseased tissues and subsequently for determining the
 CC presence or absence of cardiomyopathy or as predictor of heart disease.
 CC The present sequence is human MCIP associated protein.
 CC Note: This sequence has been described as murine splice variant in
 CC the specification, however the sequence seems to be a MCIP associated
 CC protein.
 XX Sequence 197 AA;
 SQ
 Query Match 83.3%; Score 872.5; DB 23; Length 197;
 Best Local Similarity 89.4%; Pred. No. 6.2e-89;
 Matches 168; Conservative 7; Mismatches 12; Indels 1; Gaps 1;
 QY 11 SATIACHLDPRVFDGLCRKAFESLRTYDKDTTFQYFKSRVRINFNPLSAADARL 70
 DB 11 SLLIACVANDVDFSESETRAKFESLRTYDKDTTFQYFKSRVRINFNPLSAADARL 70
 QY 71 LHKTEFLGKEMKLYFAQTLHIGSSHLAPPNDKQFLISPPASPPVGMKQVEDATPVINYD 130
 DB 71 LHKTEFLGKEMKLYFAQTLHIGSSHLAPPNDKQFLISPPASPPVGMKQVEDATPVINYD 130
 QY 131 LLYAISKLGPGEKGYELHAATDPTPSVVHVHVCSDQNEEEEEEMERMKRPKPIQTRRP 190
 DB 131 LLYAISKLGPGEKGYELHAATDPTPSVVHVHVCSDQNEEEEEEMERMKRPKPIQTRRP 190
 QY 191 EYTPIHLS 198
 DB 190 EYTPIHLS 197
 RESULT 5
 ID AAW73898
 XX AAW73898 standard; Protein; 171 AA.
 AC AAW73898;
 XX
 DT 09-APR-1999 (first entry)
 XX Human Down's Syndrome critical region 1 protein.
 DE
 XX DSCR1; Down's Syndrome critical region 1; human; pathogenic abnormality;
 KW Central Nervous System development; mental retardation; heart defect.

XX Homo sapiens.
 OS
 XX
 PN US5869318-A.
 XX
 PD 09-FEB-1999.
 XX
 PF 07-JUN-1996; 96US-0665040.
 XX
 PR 07-JUN-1995; 95ES-0001140.
 XX
 PA (PALL/) PALLEJA X E.
 XX
 PI Fuentes JI, Palleja XE, Pritchard M;
 DR WPI; 1999-152781/13.
 XX
 DR N-PSDB; AAX01282.
 XX
 PT DNA encoding foetal brain proteins - believed to be associated with
 PT Down's syndrome
 PS
 PS Claim 4: Column 15-18; 19pp; English.
 CC This sequence is encoded by the Down's Syndrome critical region 1 (DSCL1)
 CC gene of the invention. The DSCL1 gene was found to be located in the
 CC q22.1-22.2 region of human chromosome 21. An increase in the transient
 CC expression of DSCL1 mRNA in the brains of young rats, compared to
 CC expression levels in the brains of adult rats, suggests an important role
 CC for DSCL1 during the development of the Central Nervous System (CNS), and
 CC that over expression of DSCL1 may be involved in pathogenic abnormalities
 CC of mental retardation and/or heart defects as found in Down's syndrome
 CC patients.
 CC
 SO Sequence 171 AA;
 Query Match 79.7%; Score 835.5; DB 20; Length 171;
 Best Local Similarity 94.7%; Pred. No. 6,9e-85;
 Matches 160; Conservative 5; Mismatches 3; Indels 1; Gaps 1;
 QY 30 AKESLFRITDQKTPYFYSFKVRINFSPISADARLRKHTFLEKEMKLYFAQTL 89
 DB 4 AKESLFRITDQKTPYFYSFKVRINFSPISADARLRKHTFLEKEMKLYFAQTL 63
 QY 90 HIGSSHLAPNPDKQFLLSPSPVGMKQVEDATPVINYDLVLAISKLGPEKYEELHAA 149
 DB 64 HIGSSHLAPNPDKQFLLSPSPVGMKQVEDATPVINYDLVLAISKLGPEKYEELHAA 123
 QY 150 TDPSPVVHVHVCSDQNEBEEMEMKRPKPIQTRREYTPHLS 198
 DB 124 TDPSPVVHVHVCSDQNEBEEMEMKRPKPIQTRREYTPHLS 171
 RESULT 6
 AAB31788 ID AAB31788 standard; Protein; 234 AA.
 XX
 AC AAB31788;
 XX
 DT 30-APR-2001 (first entry)
 XX
 DX Amino acid sequence of a human detoxification protein.
 DE
 XX Human; detoxification protein; DETX; cancer; leukaemia; melanoma;
 KM adenocarcinoma; autoimmune disorder; inflammatory disorder;
 KM rheumatoid arthritis; asthma; atherosclerosis; autoimmune thyroiditis;
 KM psoriasis; ulcerative colitis; infection; cell proliferative disorder;
 KM actinic keratosis; arteriosclerosis; cirrhosis; hepatitis.
 OS
 OS Homo sapiens.
 XX
 FH Key location/Qualifiers
 FT Peptide 1..27
 FT /note= "signal peptide"

FT Modified-site 58
 FT /note= "potential phosphorylation site"
 FT Modified-site 75
 FT /note= "potential phosphorylation site"
 FT Modified-site 87
 FT /note= "potential phosphorylation site"
 FT Modified-site 94
 FT /note= "potential glycosylation site"
 FT Modified-site 130
 FT /note= "potential phosphorylation site"
 FT Modified-site 204
 FT /note= "potential phosphorylation site"
 FT Modified-site 215
 FT /note= "potential phosphorylation site"
 FT Modified-site 223
 FT /note= "potential phosphorylation site"
 XX
 PN W0200104305-A2.
 XX
 XX 18-JAN-2001.
 PD
 XX 06-JUL-2000; 2000WO-US18509.
 PF
 PF 07-JUL-1999; 99US-0142678.
 PR
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Tang YI, Yue H;
 XX
 DR WPI; 2001-147193/15.
 DR N-PSDB; AAF25338.
 PT New human detoxification protein and polynucleotide, useful for
 PT diagnosis, prevention and treatment of autoimmune/inflammatory
 PT disorders and cell proliferative disorders including cancer -
 XX
 XX Claim 1; Page 77; 79pp; English.
 PS
 PS The present sequence represents a human detoxification polypeptide
 CC (DETX). DETX and its (ant)agonists are useful for preventing or treating
 CC disorders associated with decreased or increased expression or activity
 CC of DETX. DETX polypeptides are useful for screening compounds that
 CC specifically binds to DETX and for identifying (ant)agonists.
 CC Diseases prevented, treated and diagnosed include cancers (e.g.
 CC leukaemia, melanoma, adenocarcinoma, cancers of the brain, bone marrow,
 CC breast, kidney, liver, pancreas, prostate and uterus),
 CC autoimmune/inflammatory disorders (e.g. rheumatoid arthritis, asthma,
 CC atherosclerosis, autoimmune thyroiditis, psoriasis and ulcerative
 CC colitis), bacterial, fungal, parasitic infections and cell
 CC proliferative disorders (e.g. actinic keratosis, arteriosclerosis,
 CC cirrhosis and hepatitis). Anti-DETX antibodies may be used as
 CC antagonists, as a targeting or delivery mechanism for bringing
 CC pharmaceutical agents into contact with cells or tissues expressing
 CC DETX and for diagnosis of DETX-related disorders.
 CC
 SO Sequence 234 AA;
 Query Match 59.2%; Score 620.5; DB 22; Length 234;
 Best Local Similarity 62.4%; Pred. No. 1.1e-60;
 Matches 121; Conservative 22; Mismatches 42; Indels 9; Gaps 2;
 QY 5 DLQDLSATLACHLDPRVVDGLCRKPFSLPTDYDKDTTPYFYSFKVRINFSPISLA 64
 DB 42 DFNDLPNSLPACVHOSVFEGBESKFKGLRITDDCVTFQFYSFKVRINFSPISLA 101
 QY 65 ADARLRKHTFLEKEMKLYFAQTLHIGSS---HLAPNPDKQFLLSPSPVGMKQV 120
 DB 102 ARARIRLHETQFGKLLKLYFAQVQRPETDQKHLAPQAKQFLLSPSPVGMKQV 161
 QY 121 EDATPVINYDLVLAISKLGPEKYEELHAA TDPSPVVHVHVCSDQNEBEEMEMKRP 180
 DB 162 NDATPVINYDLVLAISKLGPEKYEELHAA TDPSPVVHVHVCSDQNEBEEMEMKRP 216

QY 181 KPQIIOTRRPEYTP 194
 Db 217 KPQIIOTRRPGLPP 230

RESULT 7
 AAM50760
 ID AAM50760 standard; Protein; 255 AA.
 AC AAM50760;
 XX
 DT 18-APR-2002 (first entry)
 XX
 DE Human Down syndrome critical region 1-like 1 protein.
 XX
 KW Down syndrome critical region 1-like 1; DSCR1L1 alpha; human;
 KW Down syndrome; Alzheimer's disease; dementia; transgenic;
 KW neuroprotective; nontropic; anticonvulsant; diagnosis; therapy.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 11
 FT /note= "encoded by TNG"
 FT Domain 129..199
 FT /note= "nucleic acid-binding domain"
 FT Region 169..176
 FT /note= "ISPPXSP box"
 XX
 XX WO200204513-A2.
 XX
 XX 17-JAN-2002.
 XX
 XX 11-JUL-2001; 2001WO-US21982.
 XX
 XX 11-JUL-2000; 2000US-0614474.
 XX
 XX (INCY-) INCYTE GENOMICS INC.
 XX
 XX Loring JF, Tingley DW, Edwards CM, Streeter DG;
 XX WPI; 2002-164633/21.
 XX N-PSDB; ABA91457.
 XX
 XX Novel Down syndrome critical region 1-like 1 protein and nucleic acid
 XX encoding the protein useful for diagnosis and treatment of Alzheimer's
 XX disease, Down syndrome and other forms of dementia
 XX
 XX Claim 1; Page 47-49; 54pp; English.
 XX
 XX The present sequence is that of novel Down syndrome critical region
 XX 1-like 1 protein (DSCR1L1 alpha). Northern analysis indicated
 XX expression of DSCR1L1 alpha in various libraries, with the highest
 XX abundance in tissues from the nervous system, including tissues
 XX associated with schizophrenia, Huntington's disease, epilepsy and
 XX amyotrophic lateral sclerosis. An absence of DSCR1L1 expression
 XX was observed in 7 of 8 libraries from subjects with Alzheimer's
 XX disease. A claimed method for detecting differential expression of
 XX a nucleic acid encoding DSCR1L1 alpha can be used to diagnose Down
 XX syndrome, Alzheimer's disease and other forms of dementia. The
 XX purified DSCR1L1 alpha protein, or a portion of it, can be used
 XX to identify ligands which specifically bind the protein, such as
 XX a DNA or RNA molecule, peptide nucleic acid, peptide, protein,
 XX mimetic, agonist, antagonist, antibody, immunoglobulin, inhibitor
 XX or drug, for use as a therapeutic. The protein can also be used to
 XX raise antibodies useful in diagnosis and therapy.

Sequence 255 AA;
 Query Match 59.2%; Score 620.5; DB 23; Length 255;
 Best Local Similarity 62.4%; Pred. No. 1.2e-60;
 Matches 121; Conservative 22; Mismatches 42; Indels 9; Gaps 2;

QY 5 DLODPSATIACHLDPRVFDGLCRAKFESLRTYDQDTTFQYFKSKRVIRINFSPLSA 64
 Db 63 DFNDLNSLPACNVHOSVFEGEESKEKPEGLFRTYDDCVTFQLEKFRVRINFSNPKSA 122
 QY 65 ADARLRHKTEFLGKEMKLYFAQTLHIGSS- ---HLAPPNPKQFLISPPASPPVGMKOV 120
 Db 123 ARARIELHETQFRGKKLYFAQVQTPETDGDGKLHLAPPQAKQFLISPPSSPPVGMQPI 182
 QY 121 EDATPVINYDLYAIKLGEGEKYELHAATDTPSPVVHVCSQDNEEEEEEMERMKRP 180
 Db 183 NDATPVNLVLLYAVAKLGEGEKYELHAGTSTSPVVHVCDSDIEEEDPK- ---TSP 237
 QY 181 KPQIIOTRRPEYTP 194
 Db 238 KPQIIOTRRPGLPP 251

RESULT 8
 AAE18913
 ID AAE18913 standard; Protein; 197 AA.
 AC AAE18913;
 XX
 DT 17-MAY-2002 (first entry)
 XX
 DE Mouse MCIP associated proein #4.
 XX
 KW Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;
 KW heart failure; cardiomyopathy; heart disease; mouse.
 XX
 OS Mus musculus.
 XX
 XX WO200204491-A2.
 XX
 XX 17-JAN-2002.
 XX
 XX 06-JUL-2001; 2001WO-US21662.
 XX
 XX 07-JUL-2000; 2000US-216601P.
 XX 13-FEB-2001; 2001US-0782953.
 XX
 XX (TEXA) UNIV TEXAS SYSTEM.
 XX (WILL/) WILLIAMS S R.
 XX (ROTH/) ROTHERMEL B.
 XX
 XX Williams SR, Rothermel B;
 XX WPI; 2002-179698/23.
 XX N-PSDB; AAD30154.
 XX
 XX Screening for modulators of muscle calcineurin interacting protein
 XX (MCIP) binding, expression or phosphorylation, useful for treating
 XX cardiac hypertrophy or heart failure, comprises mixing MCIP,
 XX calcineurin and a test compound -
 XX
 XX Disclosure; Page 154; 174pp; English.
 XX
 XX The invention relates to muscle calcineurin interacting proteins (MCIPs)
 XX and nucleic acid molecules encoding such proteins. MCIPs form a physical
 XX complex with the catalytic subunit of calcineurin and increased levels
 XX of MCIPs correspond to a reduced ability of calcineurin to stimulate
 XX transcription of certain target genes. The invention also relates to
 XX methods for identifying modulators of MCIP binding, expression or
 XX phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin
 XX may be used for treating cardiac hypertrophy and heart failure.
 XX Antibodies to MCIP can be used in characterizing the MCIP content of
 XX healthy and diseased tissues and subsequently for determining the
 XX presence or absence of cardiomyopathy or as predictor of heart disease.
 XX The present sequence is mouse MCIP associated protein.
 XX Note: This sequence has been described as mouse MCIP2 encoding DNA in
 XX the specification, however the sequence seems to be a MCIP associated
 XX protein.

CC transcription of certain target genes. The invention also relates to
 CC methods for identifying modulators of MCIP binding, expression or
 CC phosphorylation. Inhibitors or promoters of MCIP binding, expression or
 CC may be used for treating cardiac hypertrophy and heart failure.
 CC Antibodies to MCIP can be used in characterizing the MCIP content of
 CC healthy and diseased tissues and subsequently for determining the
 CC presence or absence of cardiomyopathy or as predictor of heart disease.
 CC The present sequence is human MCIP associated protein.
 CC Note: This sequence has been described as splice variant of MCIP1
 CC initiated by exon 4 in the specification, however the sequence seems
 CC to be a MCIP associated protein.

XX Sequence 241 AA;
 Query Match 56.9%; Score 596; DB 23; Length 241;
 Best Local Similarity 61.9%; Pred. No. 6.1e-58;
 Matches 120; Conservative 20; Mismatches 44; Indels 10; Gaps 2;

QY 2 BEVDLQDLPSTAIACHLDPRVFDGLCRKAFESLFRDYKDTTFQYKSKFRVRINFSP 61
 DB 38 EMDLSDLPTSLFACSVHEAVFEAREQERFEALFTYDDQVTFQKSFRRVRINFSP 97
 QY 62 LSADARLRLHKTFLGKEMKLYFAQTLHIG-----SSHLAPPNDKQFLISPPASPPVGW 117
 DB 98 EAAARARIELHETDFNGOKLKYFAQVQMSGEVRDKSYLLPPQVQKFLISPPASPPVGW 157
 QY 118 KOVEDATPVINYDLYLAISKLGPEKVELHAATDPTPSVVHVCSQDNEEEEMER 177
 DB 158 KQSDAMPVINYDLYLAISKLGPEKVELHAGTSTPSVVHVCSQDNEEEEMER 211
 QY 178 KRPKPKIITRRPE 191
 DB 212 KNPQKTAQTRPD 225

RESULT 11
 ID AAE18918
 XX AAE18918 standard; Protein; 212 AA.
 AC AAE18918;
 XX
 DT 17-MAY-2002 (first entry)
 XX
 DE Human MCIP associated protein #5.
 KW Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;
 KW heart failure; cardiomyopathy; heart disease; human.
 XX
 OS Homo sapiens.
 XX
 PN WC200204491-A2.
 XX
 PD 17-JAN-2002.
 XX
 PF 06-JUL-2001; 2001WO-US21662.
 XX
 XX 07-JUL-2000; 2000US-216601P.
 PR 13-FEB-2001; 2001US-0782953.
 XX
 XX (TEXA) UNIV TEXAS SYSTEM.
 PA (WILL/) WILLIAMS S R.
 PA (ROTH/) ROTHERMEL B.
 XX
 XX Williams SR, Rothermel B;
 PI
 XX WPI; 2002-179698/23.
 DR N-PSDB; AAD30159.
 DR

XX Screening for modulators of muscle calcineurin interacting protein
 PT (MCIP) binding, expression or phosphorylation, useful for treating
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,
 PT calcineurin and a test compound -

PS Disclosure; Page 171-172; 174pp; English.

XX The invention relates to muscle calcineurin interacting proteins (MCIPs)
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical
 CC complex with the catalytic subunit of calcineurin and increased levels
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate
 CC transcription of certain target genes. The invention also relates to
 CC methods for identifying modulators of MCIP binding, expression or
 CC phosphorylation. Inhibitors or promoters of MCIP binding, expression or
 CC may be used for treating cardiac hypertrophy and heart failure.
 CC Antibodies to MCIP can be used in characterizing the MCIP content of
 CC healthy and diseased tissues and subsequently for determining the
 CC presence or absence of cardiomyopathy or as predictor of heart disease.
 CC The present sequence is human MCIP associated protein described in the
 CC invention.

XX Sequence 212 AA;

Query Match 56.8%; Score 595; DB 23; Length 212;
 Best Local Similarity 62.6%; Pred. No. 6.5e-58;
 Matches 119; Conservative 20; Mismatches 39; Indels 12; Gaps 2;

QY 2 BEVDLQDLPSTAIACHLDPRVFDGLCRKAFESLFRDYKDTTFQYKSKFRVRINFSP 61
 DB 19 EMDLSDLPTSLFACSVHEAVFEAREQERFEALFTYDDQVTFQKSFRRVRINFSP 78
 QY 62 LSADARLRLHKTFLGKEMKLYFAQTLHIGSSHLAPPNDKQFLISPPASPPVGW 121
 DB 79 EAAARARIELHETDFNGOKLKYFAQ-----SYLLPPQVQKFLISPPASPPVGW 132
 QY 122 DATPVINYDLYLAISKLGPEKVELHAATDPTPSVVHVCSQDNEEEEMER 181
 DB 133 DAMPVINYDLYLAISKLGPEKVELHAGTSTPSVVHVCSQDNEEEEMER 186
 QY 182 PKIITRRPE 191
 DB 187 QKIAQTRRPD 196

RESULT 12
 ID AAB58452
 XX AAB58452 standard; Protein; 142 AA.
 AC AAB58452;
 XX
 DT 14-MAR-2001 (first entry)
 XX
 DE Lung cancer associated polypeptide sequence SEQ ID 790.
 KW Human; lung cancer associated protein; neuroprotective; cytostatic;
 KW cardioactive; immunomodulatory; muscular active; vulnary;
 KW gastrointestinal; nephrotropic; anti-infective; gynecological;
 KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
 KW proliferative disorder; wound healing; infectious disease.

XX Homo sapiens.
 OS
 XX WO200055180-A2.
 PN
 XX 21-SEP-2000.
 PD
 XX 08-MAR-2000; 2000WO-US05918.
 PP
 XX 12-MAR-1999; 99US-0124270.
 PR
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA (ROSE/) ROSEN C A.
 XX Ruben SM;
 PI
 XX WPI; 2000-587514/55.
 DR N-PSDB; AAF18328.
 DR

PT Lung cancer associated gene sequences, referred to as lung cancer
PT antigens, useful for treatment, prevention, and diagnosis of disorders
PT such as lung cancer -

Claim 11: Page 1325-1326; 1425pp; English.

XX polynucleotide sequences AA17982 - AA18424 encode human lung cancer
CC associated proteins represented in AAB58106 - AAB58548. Lung cancer
CC associated proteins and polynucleotide sequences, their agonists, and
CC antagonists may have neuroprotective; cytoskeletal; cardiovascular;
CC immunomodulatory; muscular active general; vulnereary; gastrointestinal
CC general; nephrotoxic; anti-infective; gynecological; or antibacterial
CC activity. The invention also includes antibodies specific for the
CC protein or polynucleotide sequences. The lung cancer associated
CC polynucleotide sequences may be used for detection of lung cancer,
CC chromosome identification, as chromosome markers, and for numerous other
CC diagnostic or research purposes. The proteins may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders. The proteins may also be used in the treatment of wounds and
CC infectious diseases. Polynucleotide sequences AA18425 - AA18433 and
CC peptide AAB58549 are used in the course of the invention for the
CC identification and characterization of the polynucleotide and protein
CC sequences.

XX Sequence 142 AA;

Query Match 47.4%; Score 497; DB 21; Length 142;
Best Local Similarity 96.9%; Pred. No. 3.2e-47;
Matches 95; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 MEEVDIODLPATITACHLDPRVFVDGLCRAKFESLFRITYDKDTTFQYFKSKVRINFSN 60
DB 38 MEEVDIODLPATITACHLDPRVFVDGLCRAKFESLFRITYDKDTTFQYFKSKVRINFSN 97
OY 61 PLASADARLRLHKTFLGKEMKLYFAQTLLHIGSSHLAP 98
DB 98 PLASADARLRLHKTFLGKEMKLYFAQTLLHIGSSHLAP 135

RESULT 13

AA001768
ID AAG01768 standard; Protein; 111 AA.

XX AAG01768;

XX 06-OCT-2000 (first entry)

XX Human secreted protein, SEQ ID NO: 5849.

KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping.

XX Homo sapiens.

OS EPI033401-A2.

PN EPI033401-A2.

PD 06-SEP-2000.

XX 21-FEB-2000; 2000EP-0200610.

XX 26-FEB-1999; 99US-0122487.

XX (GBST) GENSET.

XX Dumas Milne Edwards J, Duclerc A, Giordano J;

DR WPI: 2000-500381/45.

XX N-PSDB; AAC01774.

PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT opating cDNAs and genomic DNAs that correspond to 5' ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -

XX Claim 13: SEQ ID 5849; 71pp + CD-ROM; English.

XX The present sequence is a polypeptide encoded by one of a large number
CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
CC were prepared from total human RNA or polyA+ RNAs derived from 30
CC different tissues. EST sequences usually correspond mainly to the 3'
CC untranslated region (UTR) of the mRNA because they are often obtained
CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
CC those cases where longer cDNA sequences have been obtained, the full 5'
CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
CC ends and can therefore be used to obtain full length cDNAs and genomic
CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
CC chromosome mapping procedures. They are used to obtain upstream
CC regulatory sequences and to design expression and secretion vectors.

XX Sequence 111 AA;

Query Match 41.9%; Score 439; DB 21; Length 111;
Best Local Similarity 84.2%; Pred. No. 6.6e-41;
Matches 85; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

OY 11 SATTACHLDPRVFVDGLCRAKFESLFRITYDKDTTFQYFKSKVRINFSNPLASADARLR 70
DB 11 SLTACVANSDFSESETRAKFESLFRITYDKDTTFQYFKSKVRINFSNPLASADARLR 70
OY 71 LKTEFLGKEMKLYFAQTLLHIGSSHLAPPNPDKQFLISPPA 111
DB 71 LKTEFLGKEMKLYFAQTLLHIGSSHLAPPNPDKQFLISPPA 111

RESULT 14

AB071467
ID ABB71467 standard; Protein; 292 AA.

XX ABB71467;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 41193.

KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.

XX Drosophila melanogaster.

OS WO200171042-A2.

PN WO200171042-A2.

PD 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI: 2001-656860/75.

XX N-PSDB; ABL15570.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -

XX Disclosure; SEQ ID NO 41193; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABU16176-ABU16175), expressed DNA
CC sequences (ABU16176-ABU16175) and the encoded proteins
CC (ABU16176-ABU16175).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ . Sequence 292 AA;

Query Match 36.2%; Score 379.5; DB 22; Length 292;
Best Local Similarity 40.7%; Pred. No. 1.2e-33;
Matches 79; Conservative 42; Mismatches 60; Indels 13; Gaps 5;

QY 3 EVD---LQDLPSTACHLDPRVVDGLCRKAKFSLRTYDKDTTFQYFKSRVRINFS 59
Db 105 EVDADSDFDLFTSIIVTNIHSEVPANPELKHMBELFTFSESATFQWLRSFRRLRVNYD 164
QY 60 NPLSAADARLRLHKTFLGKE-MKLYFAQTL-HIGSSHLAPPNDKQFLISPPASPPGVW 117
Db 165 NAIANANARIKLHGYEFNKKTVITCYPAQPTVPSNKNLQPPAPVKQFLISPPASPPAGW 224
QY 118 KQVEDATPVINYDLYLAISKLGPGKGYELHAATDPTSPVVHVCSDOENEEEMERM 177
Db 225 EPREGEPVLNHDLLAALASITPGESEHLHPQSEDQPAIIIVHTAML-----ATGPGLOV 279
QY 178 KRPKPKIIOTRPE 191
Db 280 KAP---IVQTKCPE 290

RESULT 15
ID ABB29495
XX ABB29495 standard; Peptide; 58 AA.
AC ABB29495;
XX
DT 01-FEB-2002 (first entry)
XX
DE Peptide #2146 encoded by breast cell single exon nucleic acid probe.
XX Human; microarray; single exon probe; gene expression; breast;
KW disease; cancer.
XX
OS Homo sapiens.
XX
PN WO200157271-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00662.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-496933/54.
XX
PT New spatially-addressable set of single exon nucleic acid probes,
PT useful for measuring gene expression in sample derived from human
PT breast, comprises number of single exon nucleic acid probes -
XX
PS Claim 27; SEQ ID NO 12463; 327pp + sequence listing; English.
XX
CC The invention relates to a spatially-addressable set of single exon

CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BT 474 cells. The method involves contacting
CC the probes with a collection of detectably labelled nucleic acids
CC derived from mRNA of human breast, and then measuring the label
CC bound to each probe of the microarray. The probes are useful for
CC verifying the expression of regions of genomic DNA predicted to
CC encode proteins. They are useful for gene discovery, and for
CC determining predisposition and/or prognosing breast disease. Gene
CC expression analysis is useful for assessing the toxicity of chemical
CC agents on cells. The microarray of this invention presents a far greater
CC diversity of probes for measuring gene expression, with far less bias
CC than expressed sequence tag microarrays. The method is suitable for
CC rapid production of functional information from genomic sequence. The
CC present sequence is a peptide encoded by a single exon nucleic acid
CC probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 58 AA;

Query Match 27.1%; Score 284; DB 22; Length 58;
Best Local Similarity 94.8%; Pred. No. 5.2e-24;
Matches 55; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 30 AKFESLFRITYDKDTTFQYFKSRVRINFSNPLSAADARLRLHKTFLGKEMKLYFAQ 87
Db 1 AKFESLFRITYDKDTTFQYFKSRVRINFSNPLSAADARLRLHKTFLGKEMKLYFAQ 58

Search completed: December 11, 2002, 11:42:38
Job time : 32.7184 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 11:37:42 ; Search time 24.5835 Seconds
(without alignments)
1651.161 Million cell updates/sec

Title: US-09-782-953-9

Perfect score: 1041

Sequence: 1 MPAPSMDCVSTLVACVDV.....SPKPKIOTRRPGLPPSVN 197

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SPTRMBL 21.*
2: sp_archaea.*
3: sp_bacteria.*
4: sp_fungi.*
5: sp_human.*
6: sp_invertebrate.*
7: sp_mammal.*
8: sp_mhc.*
9: sp_organelle.*
10: sp_phage.*
11: sp_plant.*
12: sp_rodent.*
13: sp_virus.*
14: sp_vertebrate.*
15: sp_unclassified.*
16: sp_rvirus.*
17: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1000	96.1	192	11 Q8VIP5	Q8vip5 mus musculus
2	926	89.0	243	11 Q8VIP4	Q8vip4 mus musculus
3	634.5	61.0	198	11 Q91WQ4	Q91wq4 mus musculus
4	588	56.5	171	4 Q96R03	Q96r03 homo sapien
5	133.5	12.8	315	3 Q9P5S9	Q9p5s9 neurospora
6	129	12.4	249	3 Q9F4A1	Q9f4a1 cryptococcus
7	95	9.1	455	11 Q920L8	Q920l8 mus musculus
8	95	9.1	535	11 Q9EQ77	Q9eq77 mus musculus
9	93	8.9	1018	6 Q28106	Q28106 bos taurus
10	92	8.8	277	3 Q13703	Q13703 schizosacch
11	91	8.7	704	5 Q17478	Q17478 hyalophora
12	90.5	8.7	1213	4 Q92923	Q92923 homo sapien
13	89	8.5	746	4 Q9BU60	Q9bu60 homo sapien
14	89	8.5	874	5 Q9VQL0	Q9vql0 drosophila
15	89	8.5	1021	4 Q15451	Q15451 homo sapien
16	89	8.5	1052	4 Q96FT1	Q96ft1 homo sapien

17	89	8.5	1251	4 Q15450	Q15450 homo sapien
18	88.5	8.5	465	4 Q96GY4	Q96gy4 homo sapien
19	88.5	8.5	501	5 Q9BKQ7	Q9bkq7 caenorhabdi
20	88.5	8.5	1130	4 Q96E12	Q96e12 homo sapien
21	88.5	8.5	1214	4 Q8TAQ2	Q8taq2 homo sapien
22	87.5	8.4	13055	5 Q09165	Q09165 caenorhabdi
23	86	8.3	1006	11 Q62901	Q62901 rattus norv
24	86	8.3	1066	4 Q8H2N5	Q8h2n5 homo sapien
25	85.5	8.2	320	5 Q95SK4	Q95sk4 drosophila
26	85.5	8.2	320	5 Q9VQR5	Q9vqr5 drosophila
27	85.5	8.2	811	2 Q59316	Q59316 clostridium
28	85.5	8.2	7576	2 Q9ZGA4	Q9zga4 streptomyce
29	85	8.2	878	5 Q9VZF8	Q9vzf8 drosophila
30	85	8.2	4025	4 Q9NR13	Q9nr13 homo sapien
31	84.5	8.1	212	10 Q9LYU3	Q9lyu3 arabidopsis
32	84.5	8.1	314	10 Q9LIV7	Q9liv7 oryza sativ
33	84.5	8.1	1114	5 Q9V180	Q9v180 drosophila
34	84.5	8.1	1222	16 Q98PR9	Q98pr9 mycoplasma
35	84.5	8.1	2701	4 Q9V520	Q9v520 homo sapien
36	84.5	8.1	3938	11 Q88778	Q88778 rattus norv
37	84.5	8.1	5002	5 Q9VXR3	Q9vvr3 drosophila
38	84	8.1	246	10 Q23143	Q23143 arabidopsis
39	84	8.1	296	10 Q8W0W9	Q8w0w9 oryza sativ
40	84	8.1	671	10 Q9FXA2	Q9fxa2 arabidopsis
41	84	8.1	1557	5 Q96652	Q96652 drosophila
42	83.5	8.0	220	10 Q9FGA3	Q9fga3 arabidopsis
43	83.5	8.0	564	5 Q9NA93	Q9na93 caenorhabdi
44	83	8.0	160	17 Q9YD46	Q9y46 aeropyrum p
45	83	8.0	643	5 Q9XUT0	Q9xut0 caenorhabdi

ALIGNMENTS

RESULT 1

ID Q8VIP5 PRELIMINARY; PRT; 192 AA.
AC Q8VIP5;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Calcineurin inhibitory protein ZAKI-4.
GN ZAKI-4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Kanou Y., Miyazaki T., Seo H., Murata Y.;
RT "calcineurin inhibitory protein ZAKI-4";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBSJ databases.
DR EMBL: AB061524; BAB71955.1; --
SQ SEQUENCE 192 AA; 21540 MW; B2DB9819F4B36AE8 CRC64;

Query Match 96.1%; Score 1000; DB 11; Length 192;
Best Local Similarity 98.4%; Pred. No. 1e-90; 2; Indels 0; Gaps 0;

Matches 189; Conservative 1; Mismatches 1; Mismatches 2; Indels 0; Gaps 0;
Qy 6 MDCDVSTLVACVVDVEVTNQEVKEKFEGLFRTYDECVTQFLFKSFRVRVINFSPKSA 65
Db 1 MDCDVSTLVACVVDVEVTNQEVKEKFEGLFRTYDECVTQFLFKSFRVRVINFSPKSA 60
Qy 66 RARIELHETQFRGKKLYFAQVQTPETDGDKLHLAPPQAKQFLISPPSPVGVKWDIS 125
Db 61 RARIELHETQFRGKKLYFAQVQTPETDGDKLHLAPPQAKQFLISPPSPVGVKWDIS 120
Qy 126 DATPVLNLDLLYAVAKLGPGEKYLHAGTSTPSVVHVCDSDMEEDDPKTSKPKEIIQ 185
Db 121 DATPVLKVDLLYAVAKLGPGEKYLHAGTSTPSVVHVCDSDMEEDDPKTSKPKEIIQ 180
Qy 186 TRRPGCLPPSVSN 197
|||||

Db 181 TRPGLPSPVSN 192

RESULT 2

Q8VIP4 PRELIMINARY; PRT; 243 AA.

AC Q8VIP4; 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Calcineurin inhibitory protein ZAKI-4 beta.
 GN ZAKI-4 BETA.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxId=10090;

RN (1)
 RP SEQUENCE FROM N.A.
 RA Kanou Y., Miyazaki T., Seo H., Murata Y.,
 RT "Calcineurin inhibitory protein ZAKI-4."
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB061525; BAB71956.1; -
 DR InterPro; IPR000504; RWA_rec_moc.
 DR Pfam; PF00076; rtm; 1.
 SQ SEQUENCE 243 AA; 27332 MW; 2CD8C1A810291851 CRC64;

Query Match 89.0%; Score 926; DB 11; Length 243;
 Best Local Similarity 94.1%; Pred. No. 2.6e-83;
 Matches 176; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Qy 11 STLAVGVDEVFPTNOEVEKEFEGLFRITYDECVTFLPKSFRRVRINFSHPKSAARARIE 70
 Db 57 NSLFACNVHQSVEEESKEKEFGLFRITYDECVTFLPKSFRRVRINFSHPKSAARARIE 116

Qy 71 LHETQFRGKKLKYFAOVQPTETDGDKLHAPPAKQFLISPPSPVGMKPISDATPV 130
 Db 117 LHETQFRGKKLKYFAOVQPTETDGDKLHAPPAKQFLISPPSPVGMKPISDATPV 176

Qy 131 LNVDLVAVAKLGGEGYELHAGTESTPSVVAVCDSDMEBEDPKTSPKIIQTRRPG 190
 Db 177 LNVDLVAVAKLGGEGYELHAGTESTPSVVAVCDSDMEBEDPKTSPKIIQTRRPG 236

Qy 191 LPPSVSN 197
 Db 237 LPPSVSN 243

RESULT 3
 Q91WQ4 PRELIMINARY; PRT; 198 AA.

AC Q91WQ4; 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Down syndrome critical region homolog 1 (human).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxId=10090;

RN (1)
 RP SEQUENCE FROM N.A.
 RA Straubeberg R.;
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC013551; AAH13551.1; -
 SQ SEQUENCE 198 AA; 22851 MW; F017C68F18ACCT87 CRC64;

Query Match 61.0%; Score 634.5; DB 11; Length 198;
 Best Local Similarity 64.4%; Pred. No. 1.1e-54;
 Matches 121; Conservative 25; Mismatches 33; Indels 9; Gaps 2;

Qy 11 STLAVGVDEVFPTNOEVEKEFEGLFRITYDECVTFLPKSFRRVRINFSHPKSAARARIE 70
 Db 57 NSLFACNVHQSVEEESKEKEFGLFRITYDECVTFLPKSFRRVRINFSHPKSAARARIE 116

Qy 71 LHETQFRGKKLKYFAOVQPTETDGDKLHAPPAKQFLISPPSPVGMKPISDATPV 130
 Db 117 LHETQFRGKKLKYFAOVQPTETDGDKLHAPPAKQFLISPPSPVGMKPISDATPV 176

Qy 131 LNVDLVAVAKLGGEGYELHAGTESTPSVVAVCDSDMEBEDPKTSPKIIQTRRPG 190
 Db 177 LNVDLVAVAKLGGEGYELHAGTESTPSVVAVCDSDMEBEDPKTSPKIIQTRRPG 236

Qy 191 LPPSVSN 197
 Db 237 LPPSVSN 243

Db 11 SSLIACVANDVSESETRAKFESLFRITYDQDTTFQYFKSFKVRINFSNPLSAAADARLR 70

Qy 71 LHETQFRGKKLKYFAOVQPTETDGDKLHAPPAKQFLISPPSPVGMKPISDATPV 130

Db 71 LHETQFRGKKLKYFAOVQPTETDGDKLHAPPAKQFLISPPSPVGMKPISDATPV 126

Qy 131 LNVDLVAVAKLGGEGYELHAGTESTPSVVAVCDSDMEBEDPKTSPKIIQTRRPG 190

Db 127 INYDLVAISLKGEGYELHAGTESTPSVVAVCDSDMEBEDPKTSPKIIQTRRPG 186

Qy 186 TRPGLPSP 193
 Db 187 TRPGLPSP 194

RESULT 4
 Q96R03 PRELIMINARY; PRT; 171 AA.

AC Q96R03; 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Down syndrome critical region protein 1.
 GN DSCR1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OC NCBI_TaxId=9606;

RN (1)
 RP SEQUENCE FROM N.A.
 RA Hua F., Wu J., Zhou Y., Zhang B., Peng X., Qiang B., Yuan J.,
 RA Qiang B.,
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF400429; AAK92478.1; -
 SQ SEQUENCE 171 AA; 19823 MW; 133001A8AEDF0BD9 CRC64;

Query Match 56.5%; Score 588; DB 4; Length 171;
 Best Local Similarity 66.5%; Pred. No. 3.6e-50;
 Matches 113; Conservative 22; Mismatches 27; Indels 8; Gaps 2;

Qy 28 VKEKEFEGLFRITYDECVTFLPKSFRRVRINFSHPKSAARARIE LHETQFRGKKLKYFAO 87
 Db 2 VYAKESLFRITYDQDTTFQYFKSFKVRINFSNPLSAAADARLR LHETQFRGKKLKYFAO 61

Qy 88 VQPTETDGDKLHAPPAKQFLISPPSPVGMKPISDATPV LNVDLVAVAKLGGEGYEL 147
 Db 62 TLHIGSS----HLAPPNDKQFLISPPSPVGMKPISDATPV INYDLVAISLKGEGYEL 117

Qy 148 YELHAGTESTPSVVAVCDSDMEBEDPKTSPKIIQTRRPG 193
 Db 118 YELHAGTESTPSVVAVCDSDMEBEDPKTSPKIIQTRRPG 167

RESULT 5
 Q9PS59 PRELIMINARY; PRT; 315 AA.

AC Q9PS59; 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE Nebula related protein.
 GN B5022.230.
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariales; Sordariaceae; Neurospora.
 OC NCBI_TaxId=5141;

RN (1)
 RP SEQUENCE FROM N.A.
 RA Schulte U., Aign V., Hohenseil J., Brandt P., Fartmann B., Holland R.,
 RA Nyakatura G., Mewes H.W., Mannhaupt G.,
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 RN (2)
 RP SEQUENCE FROM N.A.

Qy 11 STLAVGVDEVFPTNOEVEKEFEGLFRITYDECVTFLPKSFRRVRINFSHPKSAARARIE 70
 Db 57 NSLFACNVHQSVEEESKEKEFGLFRITYDECVTFLPKSFRRVRINFSHPKSAARARIE 116

Qy 71 LHETQFRGKKLKYFAOVQPTETDGDKLHAPPAKQFLISPPSPVGMKPISDATPV 130
 Db 117 LHETQFRGKKLKYFAOVQPTETDGDKLHAPPAKQFLISPPSPVGMKPISDATPV 176

Qy 131 LNVDLVAVAKLGGEGYELHAGTESTPSVVAVCDSDMEBEDPKTSPKIIQTRRPG 190
 Db 177 LNVDLVAVAKLGGEGYELHAGTESTPSVVAVCDSDMEBEDPKTSPKIIQTRRPG 236

Qy 191 LPPSVSN 197
 Db 237 LPPSVSN 243

```

RA German Neurospora genome project;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL355932; CAB91442.1; -.
SQ SEQUENCE 315 AA; 34194 MW; 76792987653C3AB7 CRC64;

  Query Match      12.8%; Score 133.5; DB 3; Length 315;
  Best Local Similarity 37.1%; Pred. No. 4.9e-05;
  Matches 36; Conservative 12; Mismatches 46; Indels 3; Gaps 2;

QY 49 KSPRRVRINFSHPKSAARAIHELHETQFRGKGLKLYFAQVQTPETDGDKLHLAPPOPAKQ 108
Db 101 KSPRRIRIVTFDEQAIAVRSVWDGEAILGERCRVVFQOPTIDVSAADKHLALPDAGKL 160
QY 109 FLISPPSSPPGVGKHP-ISDA--TPVLNLDLYAVAKL 142
Db 161 FFISPPSPSHDWEQMEDAPMTVHAEADLAELAKL 197

RESULT 6
Q9P4A1
ID Q9P4A1 PRELIMINARY; PRT; 249 AA.
AC Q9P4A1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Calcineurin-binding protein.
GN CBP1.
OS Cryptococcus neoformans var. neoformans.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
OX NCBI_TaxID=40410;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H99;
RX Goriach J., Fox D.S., Cutler N.S., Cox G.M., Perfect J.R., Heitman J.;
RA "Identification and characterization of a highly conserved calcineurin
RT binding protein, CBP/calciopressin, in Cryptococcus neoformans.";
RL EMBO J. 19:3618-3629(2000).
DR EMBL; AF230799; AAF63734.1; -.
SQ SEQUENCE 249 AA; 27293 MW; 7B757183FA5EBF09 CRC64;

  Query Match      12.4%; Score 129; DB 3; Length 249;
  Best Local Similarity 34.0%; Pred. No. 0.0001;
  Matches 35; Conservative 15; Mismatches 35; Indels 18; Gaps 3;

QY 81 LKLYFAQVQTPETDGDKLHLAPPOPAKQLISPPSPVGVGKPISDATP---VLNVDLLY 137
Db 107 LRLHLYPPTLNPDPATTHLAPPPLPHNPLISPPGSPGWEPAEAPNRIILPEDLQR 166
QY 138 AVAKLPGGKYEHLHAGTESTP-----SVVVHVCDSDMBE 171
Db 167 AL-----ETLEINSGSKADGKEIILDEGGVRVQVEDTTKQE 203

RESULT 7
Q920L8
ID Q920L8 PRELIMINARY; PRT; 455 AA.
AC Q920L8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Fcα/m receptor (Fragment).
GN FcαR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Shimizu Y., Honda S., Yotsumoto K., Tahara-Hanaoka S., Eyre H.J.,
RA Sutherland G.D., Endo Y., Shibuya K., Koyama A., Nakauchi H.,
RA Shibuya A.;

```

```

RT "Fcα/m receptor is a single gene-family member closely related to
RT polymeric immunoglobulin receptor on chromosome 1.";
RL Immunogenetics 0:0-0(2001).
DR EMBL; AB071978; BAB71750.1; -.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003596; IG_V.
DR InterPro; IPR000508; SigPbase.
DR Pfam; PF00047; IG; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00501; SPASE_1_1; UNKNOWN_1.
KW Receptor.
FT NON_TER
SQ SEQUENCE 455 AA; 48810 MW; 628913C33A4AC365 CRC64;

  Query Match      9.1%; Score 95; DB 11; Length 455;
  Best Local Similarity 25.0%; Pred. No. 0.48;
  Matches 29; Conservative 19; Mismatches 50; Indels 18; Gaps 4;

QY 87 QVQTPETDGDK-----LHLAPPOPAKQLISPPS---SPVGVGKPISDATPVLNVDLLYAV 139
Db 243 QGTTPTDGPREDTDRVSPAPRKTGTTRPSALISEHVTWETLQDKTEVSKQQLHSL 302
QY 140 AKLGPGEKYE-----LHAGTESPSVVHVCDSDMBEEDPKTSPKPKIQTTRP 189
Db 303 BELSPAPSAQTLNATCLEVASEGRSI-----DGSLENTTESSPPTPSQLSVAGP 353

RESULT 8
Q9EQ7
ID Q9EQ7 PRELIMINARY; PRT; 535 AA.
AC Q9EQ7;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Fcα/m receptor.
GN FCαR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21170225; PubMed=11062505;
RA Shibuya A., Sakamoto N., Shimizu Y., Shibuya K., Osawa M.,
RA Hironaka T., Eyre H.J., Sutherland G.R., Endo Y., Fujita T.,
RA Miyabayashi T., Sakano S., Tsuji T., Nakayama E., Phillips J.H.,
RA Lanier L.L., Nakauchi H.;
RT "Fcα/m receptor mediates endocytosis of IgM-coated microbe.";
RL Nat. Immunol. 1:441-446(2000).
DR EMBL; AB048834; BAB17312.1; -.
DR MGD; MGI:1927803; FcαR.
DR InterPro; IPR003599; IG.
DR InterPro; IPR003600; IG_Like.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR000508; SigPbase.
DR Pfam; PF00047; IG; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00410; IG_Like; 1.
DR PROSITE; PS00501; SPASE_1_1; UNKNOWN_1.
KW Receptor.
SQ SEQUENCE 535 AA; 57696 MW; B275B7C70151C75 CRC64;

  Query Match      9.1%; Score 95; DB 11; Length 535;
  Best Local Similarity 25.0%; Pred. No. 0.59;
  Matches 29; Conservative 19; Mismatches 50; Indels 18; Gaps 4;

QY 87 QVQTPETDGDK-----LHLAPPOPAKQLISPPS---SPVGVGKPISDATPVLNVDLLYAV 139
Db 323 QGTTPTDGPREDTDRVSPAPRKTGTTRPSALISEHVTWETLQDKTEVSKQQLHSL 382
QY 140 AKLGPGEKYE-----LHAGTESPSVVHVCDSDMBEEDPKTSPKPKIQTTRP 189
Db 383 BELSPAPSAQTLNATCLEVASEGRSI-----DGSLENTTESSPPTPSQLSVAGP 433

```

RESULT 9

ID 028106 PRELIMINARY; PRT; 1018 AA.

AC 028106; PRELIMINARY; PRT; 1018 AA.
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE F3/11/contactin precursor.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Kuminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OC NCBI_TaxId=9913;
 RN [1]

RP TISSUE=FROM N.A.

RC MEDLINE=95369697; PubMed=7642103;
 RA Watanabe K., Shimazaki K., Hosoya H., Fukumachi F., Takenawa T.;
 RT "Cloning of the DNA encoding neural adhesion molecule F3 from bovine
 brain";
 RL Gene 160:245-248 (1995).

DR EMBL; D32135; BAA06861.1; -.
 DR HSSP; P08921; 1A7B.
 DR InterPro; IPR003961; FN_III.
 DR InterPro; IPR003598; IG_C2.
 DR InterPro; IPR003600; IG_1like.
 DR InterPro; IPR003006; IG_MHC.
 DR InterPro; IPR002052; N6_Mtase.
 DR Pfam; PF00041; Im3; 4.
 DR Pfam; PF00047; Ig; 6.
 DR SMART; SMO0060; FN3; 4.
 DR SMART; SMO0408; IGc2; 4.
 DR SMART; SMO0410; IG_1like; 2.
 DR PROSITE; PS00092; N6_MTASE; UNKNOWN_1.
 DR Immunoglobulin domain; Signal.
 KM SIGNAL
 FT SIGNAL 1 20
 SQ SEQUENCE 1018 AA; 113384 MW; 551FC10ED7920341 CRC64;

Query Match 8.9%; Score 93; DB 6; Length 1018;

Best Local Similarity 21.4%; Pred. No. 2.1; Matches 70; Indels 36; Gaps 7;

QY 9 DVSTLVACVDEVEFTNOEYKEKEFGLEFRTYDECVTFOLEKSFRRVYRINFSHPKSNARAR 68
 DB 806 DAPSEAPTVAGVAVLSSSEISVMEH-----VVEKIVESY-QIRYASSHKKAAR 856
 QY 69 IELEHTQFRGKKLKL-----YFAQVOTPETDGDKHLAPPOPAKQFLI--SPSSSP--- 118
 DB 857 VQVASQEGYARLEMLLPDIOYFVEVRACNSAG---CGPPSDMTETFTKKAPSPRPRII 912
 QY 119 -----VGWK---PISDAFVNIYDLYAVAKLGGSEKELHAGTESTP 158
 DB 913 SSVRSGRYITIMDHVVALSNSTSTVTGKVLKRPDGDHDKLSTYTKHSLIEVP 965

RESULT 10

ID 013703 PRELIMINARY; PRT; 277 AA.

AC 013703; PRELIMINARY; PRT; 277 AA.
 DT 01-JAN-1999 (TREMBlrel. 09, Created)
 DT 01-JAN-1999 (TREMBlrel. 09, Last sequence update)
 DT 01-JAN-1999 (TREMBlrel. 09, Last annotation update)
 DE Hypothetical 29.6 kDa protein C13f5.04c in chromosome I.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OC NCBI_TaxId=4996;
 RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=972;
 RA Brown D., Churcher C.M., Barrell B.G., Rajandream M.A., Wood V.;
 RA Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z99091; CAB1767.1; -.
 KW Hypothetical protein.
 FT DOMAIN 267 270
 SQ SEQUENCE 277 AA; 29622 MW; 540CF08359A06546 CRC64;

Query Match 8.8%; Score 92; DB 3; Length 277;

Best Local Similarity 22.0%; Pred. No. 0.52; Matches 45; Conservative 24; Mismatches 76; Indels 60; Gaps 7;

QY 42 CYTPOLEFSFRVRINP-----SHPSAR-----ARILEHTQFRGK 79
 DB 2 CFFRLVSSFAILRLISFLSIKSIKCKRALQVAMPQKAKTPSNHATELQOSTNST 61
 QY 80 KLUKLFACVOTPETDGDKHLAP--OPAKQFLISPPSPVGMKP-----ISDA 127
 DB 62 TLPTGEAAVETNASHETSPALPTSPASLSISPTKSAVSSBPVADVKSLSTPA 121
 QY 128 TPVLNVDLYAVAKLGPGEKY-----LHAGTSTPSV-----VHVCDSDME 171
 DB 122 APQLN-----SPSHSYETTPSTSTITENLPTIDPTRSRSSSHIQSLSPESKO 171
 QY 172 EEDPRTSPKPIQTRRPGLPSPVS 196
 DB 172 TSDGHRPPSPSTSTSTSTIDPSVA 196

RESULT 11

ID 017478 PRELIMINARY; PRT; 704 AA.

AC 017478; PRELIMINARY; PRT; 704 AA.
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Arylphorin precursor.
 OS Arylphorin cecropia (Cecropia moth).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 OC Bombycoidea; Saturniidae; Saturniinae; Attacini; Hyalophora.
 OC NCBI_TaxId=7123;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=FAT BODY;
 RA Massey H.C. Jr.;
 RT "The Evolution of Insect Storage Proteins: Driven by Composition and
 RT Constrained by Sequence";
 RL Theis (1995), University of Pennsylvania.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=FAT BODY;
 RX MEDLINE=98329402; PubMed=9664700;
 RA Bumester T., Massey H.C. Jr., Zakharin S.O., Benes H.;
 RT "The evolution of hexamerins and the phylogeny of insects.";
 DR J. Mol. Evol. 47:93-108 (1998).
 DR EMBL; AF032396; AAB86644.1; -.
 DR HSSP; P04253; 1ILA.
 DR InterPro; IPR000896; Hemocyanin.
 DR InterPro; IPR005203; hemocyanin_C.
 DR InterPro; IPR005204; hemocyanin_N.
 DR Pfam; PF003722; hemocyanin_1.
 DR Pfam; PF03723; hemocyanin_C; 1.
 DR Pfam; PF03722; hemocyanin_N; 1.
 DR PRINTS; PR00187; HAEMOCYANIN.
 DR PROSITE; PS00210; HEMOCYANIN_2; 1.
 KM Signal.
 FT SIGNAL 1 16
 FT CHAIN 17 704
 SQ SEQUENCE 704 AA; 83788 MW; C1805EDA6ACD4626 CRC64;

Query Match 8.7%; Score 91; DB 5; Length 704;

Best Local Similarity 27.0%; Pred. No. 2.1;

```
Matches 40; Conservative 24; Mismatches 56; Indels 28; Gaps 6;
QY 7 DCDVSTLVACVVDVEFTNQEVKFKGLPRT-----YDECVTFO-----LKFSPR 52
Db 64 DYDVE-----ANIDYTNKKAVEFLKLYRTGYLPKYFEFSIFPHKLRBEAIALPHFY 117
QY 53 RVRINFSPKSAARIELHETOPRGKKLKYFAQVOTPETDGDKLHLAPQPAQFLIS 112
Db 118 YAKDFETFKSAAPARVHLNEGQP---LYAYYIAVIQNRDTHGVLDP-APYEVYVPPQFVN 173
QY 113 PPSPPPVGMKPIDA-----TPVLNYDLL 136
Db 174 MDTTVRIYRTKMDGILHPTKAINYGIV 201

RESULT 12
Q92923
ID Q92923 PRELIMINARY; PRT; 1213 AA.
AC Q92923;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE SWI/SNF complex 170 kDa subunit.
GN BAF170.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96397413; PubMed=8804307;
RA Wang W., Xue Y., Zhou S., Kuo A., Cairns B.R., Crabtree G.R.;
RT "Diversity and specialization of mammalian SWI/SNF complexes.";
RL Genes Dev. 10:2117-2130(1996).
[2]
RP SEQUENCE FROM N.A.
RA Wang W., Cote J., Xue Y., Zhou S., Khavari P.A., Biggar S.R.,
RA Murchard C., Kaipana G.V., Goff S.P., Yaniv M., Workman J.L.,
RA Crabtree G.R.;
RT "Purification and biochemical heterogeneity of the mammalian SWI-SNF
RT complex.";
RL EMBO J. 15:0-0(1996).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 MYB-LIKE DOMAIN.
DR EMBL; U66616; AAC50694.1; -.
DR InterPro: IPR001357; BRCT.
DR InterPro: IPR000953; Chromo.
DR InterPro: IPR001005; Myb_DNA_binding.
DR Pfam; PF00533; BRCT; 1.
DR Pfam; PF00249; myb_DNA-binding; 1.
DR SMART; SM00298; CHROMO; 1.
DR SMART; SM00395; SANT; 1.
DR PROSITE; PS50090; MYB 3; 1.
KW DNA-binding; Nuclear protein.
SQ SEQUENCE 1213 AA; 132732 MW; D2139FD7E7CB5235 CRC64;

Query Match
Best Local Similarity 21.0%; Pred. No. 4.5;
Matches 43; Conservative 26; Mismatches 69; Indels 67; Gaps 9;
QY 10 VSTLVACVVDVEFTNQEVKFKGLPRTYDECVTTFQFKSPRRVINFSPKSAARARI 69
Db 884 IKSLVALLVETOM-----KLEIKLRHFELETIM-----DREREAL 920
QY 70 ELHETOPRGKKLKYFAQVOTPETDGDKLHL-----APPQAKQFLISPPSPVGVWK 122
Db 921 EYQEQQLADRQAFHNEQLKYPWRARQHQFQMHQOQQQPPPA-----LPPGSQPI--P 973
QY 123 PISDATPVLYNLLLYAVA-----KLGRCEKVELHAGTESTPVSVVHVCDSD 168
Db 974 PTCAAGPPAVHGLAVPASVVPAPAGSGAPPGSLGRSEIQIG-QAGSTRGP----- 1022
QY 169 MESEEDPKTSKPKKIOTRRPGLPP 193
```

```
Db 1023 --QQQQPAGAPQPGAVP---PGVPP 1042

RESULT 13
Q9BU60
ID Q9BU60 PRELIMINARY; PRT; 746 AA.
AC Q9BU60;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical 79.2 kDa protein (fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC002875; AA02875.1; -.
DR InterPro: IPR002965; P-rich_extensn.
DR PRINTS; PRO1217; PRICHEXTENSN.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 746 AA; 79187 MW; 35A186C5A7ADEA8B CRC64;

Query Match
Best Local Similarity 8.5%; Score 89; DB 4; Length 746;
Matches 29; Conservative 13; Mismatches 44; Indels 12; Gaps 5;
QY 101 APPQPAKQLISPPSPVGMKPIDATPVLYNVDLLYAVAKLPGKVELHAGTESTPVS 160
Db 592 APPPTLPALPPSPSP-KVQPEPEP-----GLLEVEE--PGTEEE--RGADTAPIL 642
QY 161 VHVCDSDME---EEEDPKTSKPKKIOTRRPGLPPSV 195
Db 643 APEALPSQGEVERGESPAAGPPQELVEEPSAPPTL 680

RESULT 14
Q9VQLO
ID Q9VQLO PRELIMINARY; PRT; 874 AA.
AC Q9VQLO;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE CGI7258 protein.
GN CGI7258.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abiri J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.V., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
```

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fietemann W.,
 RA Foeller C., Gabrielsen A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jatali M., Katush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kratz C., Kravitz S., Kulp D., Lai Z.,
 RA Laeko P., Lei Y., Lavitky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mettel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Mekulov G., Milhina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shie B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spreading A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tecor C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang Q., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003581; AAF51156.1; -
 DR FLYBase: FBgn0031496; CG1258.
 SQ SEQUENCE 874 AA; 103694 MW; 5F56D1C0E7A01D9A CRC64;

Query Match 8.5%; Score 89; DB 5; Length 874;
 Best Local Similarity 21.0%; Pred. No. 42;
 Matches 50; Conservative 28; Mismatches 74; Indels 86; Gaps 9;

OY 6 MDCDVSTLVACVDFVFTNOEYKEKEGL-FTYDECYTFOLFKS-FRRVRI----- 56
 DB 201 MDCWTEQGAQVEIPAFIEQV-ISFDALRYNFTKQVQADESKALRCRTPSPLTST 259
 OY 57 -NPSHKSAAARARIELHETQFRGK-----LK----- 82
 DB 260 EEFRRPMRAAAROMLHVTRHEEDLQENKQAOQPEGEPSLSKRESSTCSTSDGIDS 319
 OY 83 -----LYFAQVOTPPTDDGKHLARPQAK--QPLISPPSPVGMKPRISDAPVL 131
 DB 320 VTSDLAEALYELQLEAAKENEELDKELVPPKLVQFQLPKKSPTLTTHFVGD----- 375
 OY 132 NYDLVAVAKLGGEKELHAGTSTPSVYVHVCSDMEEDDPKTPSPKXIQTTRP 189
 DB 376 -----EAEPEQEPILIRI-----ELPKVYVAKPEBELKQF 406

RESULT 15

O15451 PRELIMINARY; PRT; 1021 AA.
 AC O15451;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Proline and glutamic acid rich nuclear protein isoform
 DE (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID:9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Vadlamudi R.K., Chung J.K., Shin J.;
 RT "p160.2 an isoform of proline and glutamic acid rich nuclear protein
 RT p160.";
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U88154; AAC17709.1; -
 DR InterPro: IPR002965; P rich extenam.
 DR PRINTS: PRO1217; PRICHEXTEN5N.
 KW Nuclear protein.
 FT NON_TER 1 1

SQ SEQUENCE 1021 AA; 109136 MW; C8951B4D67698D3 CRC64;

Query Match 8.5%; Score 89; DB 4; Length 1021;
 Best Local Similarity 29.6%; Pred. No. 5.1;
 Matches 29; Conservative 13; Mismatches 44; Indels 12; Gaps 5;

OY 101 APPQAKFLISPPSSPPVGMKPRISDAPVUNYDLVAVAKLGGEKELHAGTSTPSV 160
 DB 867 APPPTLPALPPSPSP-KVQPEPEEP-----GILLVEE--PGTEBE--RGADTAFTL 917
 OY 161 VHVCDSDME--EEDPKTSPKXIQTTRRGLPPSV 195
 DB 918 APBALPSQGEYERBGESPAAGPPQELVVEEPPSPPTL 955

Search completed: December 11, 2002, 11:41:09
 Job time : 27.5835 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 11:37:42 ; Search time 7.3086 Seconds
(without alignments)
1117.976 Million cell updates/sec

Title: US-09-782-953-9

Perfect score: 1041

Sequence: 1 MPAPSMDCDSTLVACVDV.....SPKPIIQTTRPGLPPSVSN 197

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1041	100.0	197	1	CCP2 MOUSE
2	1007	96.7	197	1	CCP2 HUMAN
3	663.5	63.7	239	1	CCP3 MOUSE
4	658.5	63.3	241	1	CCP3 HUMAN
5	641	61.6	197	1	CCP1 CRIGR
6	622	59.8	197	1	CCP1 HUMAN
7	615.5	59.1	198	1	CCP1 MOUSE
8	404	38.8	292	1	NLA DROME
9	322.5	31.0	207	1	CCP1 CAEL
10	107.5	10.3	163	1	YAP1 SCHPO
11	101	9.7	454	1	YAP1 HUMAN
12	90	8.6	526	1	CAP YEAST
13	87.5	8.4	211	1	RCN1 YEAST
14	84.5	8.1	1827	1	MAP2 HUMAN
15	83	8.0	1206	1	FM14 MOUSE
16	83	8.0	1232	1	YOUS CAEL
17	83	8.0	1468	1	FMN1 MOUSE
18	82.5	7.9	1377	1	NEO1 RAT
19	82	7.9	2464	1	MAPB MOUSE
20	82	7.9	3703	1	ABF1 HUMAN
21	81.5	7.8	3726	1	ABF1 MOUSE
22	81	7.8	559	1	ECM1 MOUSE
23	81	7.8	2459	1	MAPB RAT
24	79	7.6	909	1	CNG4 HUMAN
25	78.5	7.5	497	1	HMS5 DROME
26	78.5	7.5	836	1	NOT3 YEAST
27	78	7.5	1018	1	CONT HUMAN
28	78	7.5	2468	1	MAPB HUMAN
29	77.5	7.4	508	1	V56K PLRV1
30	77.5	7.4	928	1	CHS2 EXODE
31	77.5	7.4	1461	1	NEO1 HUMAN
32	77	7.4	834	1	V2A PSVJ
33	77	7.4	1258	1	GLI2 HUMAN

34 77 7.4 1302 1 NRG DROME
35 76.5 7.3 397 1 A2BP HUMAN
36 76.5 7.3 452 1 ODO2 SCHPO
37 76.5 7.3 1733 1 VNUA_PRVA
38 76 7.3 551 1 CAP SCHPO
39 76 7.3 728 1 P85B HUMAN
40 76 7.3 793 1 S3A1 HUMAN
41 76 7.3 857 1 NFM CHICK
42 76 7.3 870 1 Y563 HUMAN
43 76 7.3 930 1 SM6C HUMAN
44 76 7.3 1493 1 NEO1 MOUSE
45 75.5 7.3 614 1 ACES_RAT

ALIGNMENTS

RESULT 1

ID CCP2 MOUSE STANDARD; PRT; 197 AA.
AC Q9JHG2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Calcipressin 2 (Down syndrome candidate region 1-like protein 1)
DE (Myocyte-enriched calcineurin interacting protein 2) (MCIP2).
GN DSCR1L1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20187590; PubMed=10722714;
RA Rothermel B., Vega R.B., Yang J., Wu H., Bassel-Duby R.,
RA Williams R.S.;
RT "A protein encoded within the Down syndrome critical region is
RT enriched in striated muscles and inhibits calcineurin signaling.";
RL J. Biol. Chem. 275:8719-8725(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BA1B/c; TISSUE=Brain;
RX MEDLINE=20534792; PubMed=11080588;
RA Striappoli P., Petrini M., Lenzi L., Carinci P., Zannotti M.;
RT "The murine DSCR1-like (Down syndrome candidate region 1) gene family:
RT conserved synteny with the human orthologous genes.";
RL Gene 257:223-232(2000).
CC -!- FUNCTION: Inhibits calcineurin-dependent transcriptional responses
CC by binding to the catalytic domain of calcineurin A. Could play a
CC role during central nervous system development.
CC -!- TISSUE SPECIFICITY: Highest expression in heart, skeletal muscle
CC and brain. Lower expression in all other tissues.
CC -!- SIMILARITY: BELONGS TO THE DSCR1 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF237791; AAF63487.1; -;
CC EMBL; AF237887; AAF62538.1; -;
CC MGD; MGI:1858219; Dscr1l1.
SQ SEQUENCE 197 AA; 22025 MW; CE1306B7B03E70F1 CRC64;

Query Match 100.0%; Score 1041; DB 1; Length 197;
Best Local Similarity 100.0%; Pred. No. 1.le-83;
Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MPAPSMDCDSTLVACVDVVEFTNQVKPEGLFRITYDECVTFQLPKSFRRVINFSSH 60
|||||

```

Db      1 MPAPSMDCVSTLVACVDEVEFTNOEVEKEBGLFRITVDECVTQOLFSPRRVAINSH 60
Qy      61 PSAAARARIELEHETQFRGKKLKYFAQVQTPETDGDKLHAPPAKQPLISPPSPPVG 120
Db      61 PSAAARARIELEHETQFRGKKLKYFAQVQTPETDGDKLHAPPAKQPLISPPSPPVG 120
Qy      121 WKPISDATPVLYNLDLYAVAKLGEGEYELHAGTESTPSVVHVCDSDMEBEDPKTSPK 180
Db      121 WKPISDATPVLYNLDLYAVAKLGEGEYELHAGTESTPSVVHVCDSDMEBEDPKTSPK 180
Qy      181 PKIOTRRRGLPPSVSN 197
Db      181 PKIOTRRRGLPPSVSN 197

RESULT 2
CCP2_HUMAN      STANDARD;      PRT;      197 AA.
ID      CCP2_HUMAN
AC      Q14206;
DT      30-MAY-2000 (Rel. 39, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Calcipressin 2 (thyroid hormone-responsive protein ZAKI-4) (Down
DE      syndrome candidate region 1-like 1) (Myocyte-enriched calcineurin
DE      interacting protein 2) (MCIP2).
GN      DSCR1L1 OR ZAKI4.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX      Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Fibroblast;
RX      MEDLINE=96278928; PubMed=8662924;
RA      Miyazaki T., Kanou Y., Murata Y., Ohmori S., Niwa T., Maeda K.,
RA      Yamazaki H., Seo H.;
RT      "Molecular cloning of a novel thyroid hormone-responsive gene, ZAKI-4,
RT      in human skin fibroblasts."
RT      J. Biol. Chem. 271:14567-14571(1996).
CC      - FUNCTION: Inhibits calcineurin-dependent transcriptional responses
CC      by binding to the catalytic domain of calcineurin A. Could play a
CC      role during central nervous system development.
CC      - TISSUE SPECIFICITY: Expressed in fibroblasts, heart, brain, liver,
CC      and skeletal muscle but not in placenta, lung, kidney and
CC      pancreas.
CC      - INDUCTION: By thyroid hormone.
CC      - SIMILARITY: BELONGS TO THE DSCR1 FAMILY.
CC      - This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.emb-eb.ch/announce/
CC      or send an email to license@eb-eb.ch).
CC      EMBL: DB3407; BAA11911.1; ALT_INIT.
CC      DR      Genew: HGNC:3041; DSCR1L1.
CC      DR      MIM: 604876;
SQ      SEQUENCE 197 AA; 21955 MW; AF39735F6661C8E0 CRC64;

Query Match      96.7%; Score 1007; DB 1; Length 197;
Best Local Similarity 96.4%; Pred. No. 1e-80;
Matches 190; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy      121 WKPISDATPVLYNLDLYAVAKLGEGEYELHAGTESTPSVVHVCDSDMEBEDPKTSPK 180
Db      121 WKPISDATPVLYNLDLYAVAKLGEGEYELHAGTESTPSVVHVCDSDMEBEDPKTSPK 180
Qy      181 PKIOTRRRGLPPSVSN 197
Db      181 PKIOTRRRGLPPSVSN 197

RESULT 3
CCP3_MOUSE      STANDARD;      PRT;      239 AA.
ID      CCP3_MOUSE
AC      Q9UKK0; Q9CXB7;
DT      16-OCT-2001 (Rel. 40, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Calcipressin 3 (Down syndrome candidate region 1-like protein 2)
DE      (Myocyte-enriched calcineurin interacting protein 3) (MCIP3).
GN      DSCR1L2.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      Mus musculus.
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=BA1B/c; TISSUE=Brain;
RA      Striipoli P., Petrini M., Lenzi P., Zannotti M.;
RT      "The murine DSCR1-like gene family."
RL      Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN      [2]
RP      SEQUENCE OF 40-239 FROM N.A.
RC      STRAIN=C57BL/6J; TISSUE=Embryonic head;
RX      MEDLINE=21085660; PubMed=11217851;
RA      Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA      Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA      Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA      Saito T., Okazaki Y., Gojodori T., Bono H., Kasukawa T., Saito R.,
RA      Kadota K., Matsuda H.A., Ashburner M., Batilov S., Casavant T.,
RA      Pilschmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
RA      Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA      Schirral L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA      Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Baren G.,
RA      Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA      Brownstein S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA      Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA      Lyons P., Marchionni L., Mashima J., Mazzei J., Mombaerts P.,
RA      Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA      Saeki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA      Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilmink L.,
RA      Wyszewski-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohetsuki S.,
RA      Hayashizaki Y.;
RT      "Functional annotation of a full-length mouse cDNA collection."
RL      Nature 409:685-690(2001).
CC      - FUNCTION: Inhibits calcineurin-dependent transcriptional responses
CC      by binding to the catalytic domain of calcineurin A. Could play a
CC      role during central nervous system development (by similarity).
CC      - SIMILARITY: BELONGS TO THE DSCR1 FAMILY.
CC      - This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.emb-eb.ch/announce/
CC      or send an email to license@eb-eb.ch).
CC      EMBL: AF237888; AAF62539.1; -
CC      DR      EMBL: AK019377; BAB31687.1; -
CC      DR      MGD: MGI:1858220; Dscr1l2.
SQ      SEQUENCE 239 AA; 27153 MW; 1B2687B47B84D272 CRC64;

Query Match      63.7%; Score 663.5; DB 1; Length 239;
Best Local Similarity 72.0%; Pred. No. 9.3e-51;

```

	Query Match	63.3%;	Score 658.5;	DB 1;	Length 241;
	Best Local Similarity	70.1%;	Pred. No. 2.5e-50;		
	Matches 131;	Conservative 17;	Mismatches 38;	Indels 1;	Gaps 1;
Qy	11	STLVACVVDVEVTNORVKBFEGELPRTYDECVTFQLFKGFRRVINFSHPKSAARAE	70	:	:
Dd	47	TSLFACSVEHVAEFAEROKERFEALFTIYDQTFLFKSFRVRVINFSKEAARAARIE	106	:	:
Qy	71	LHETQRGGKLKYFAQOVPTDTGDGLHLAPQPAQKFLISPPSSPVPVGWKPISDAATPV	130	:	:
Dd	107	LHETDFNGQKLKYFAQOVMSGEVRDKSYLLPPQPVPKQLISPSPASPPWGKQSEDAMPV	166	:	:
Qy	131	LNVDLLAVAKLGPGEKEYELHAGTESTPSVVVHVCDSDMEEEDPKTSRPKKLIQTTRPG	190	:	:
Dd	167	INTDLLCAVLKGPEKEYELHAGTESIPSVVVHVCESETTEEBBTK-NPKKIAQITRPD	225	:	:
Qy	191	LPPSVSN	197	:	:
Dd	226	pptAAALN	232	:	:

[illegible]

```

DB 127 INVDLYIAISKLGPEKXELHATDTTPSVVHVCSDEQENEEEMERKPKXIQT 186
QY 187 RRRGLPP 193
DB 187 RREYTP 193

RESULT 6
CCPI_HUMAN
ID CCPI_HUMAN STANDARD; PRT; 197 AA.
AC P53805; O00583; O00583; Q9UFI5; Q9BU69;
DT 01-OCT-1996 (Rel. 34, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Calcipressin 1 (Down syndrome critical region protein 1) (Myocyte-
enriched calcineurin interacting protein 1) (MCIP1) (Adapt78).
GN DSCR1 OR DSCR1 OR ADAPT78.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 3 AND 4).
RX MEDLINE=96121593; PubMed=8555418;
RA Fuentes J.-J., Pritchard M.A., Planas A.M., Bosch A., Ferrer I.,
RA Estivill X.;
RT "A new human gene from the Down syndrome critical region encodes a
RT protein-rich protein highly expressed in fetal brain and heart.";
RL Hum. Mol. Genet. 4:1935-1944 (1995).
RN [2]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=97468152; PubMed=9325060;
RA Fuentes J.-J., Pritchard M.A., Estivill X.;
RT "Genomic organization, alternative splicing, and expression patterns
RT of the DSCR1 (Down syndrome candidate region 1) gene.";
RL Genomics 44:358-361 (1997).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX TISSUE=Lung;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/Genbank/DBJ databases.
RN [4]
RP SEQUENCE OF 1-164 FROM N.A. (ISOFORM 2).
RX TISSUE=Breast;
RA Crawford D.R., Leahy K.P., Davies K.J.A.;
RT "Adapt78, a calcium and oxidant-inducible RNA.";
RL Submitted (Apr-1996) to the EMBL/Genbank/DBJ databases.
RN [5]
RP CHARACTERIZATION.
RX MEDLINE=20320698; PubMed=10861295;
RA Fuentes J.-J., Genesca L., Kingsbury T.J., Cunningham K.W.,
RA Perez-Riba M., Estivill X., de la Luna S.;
RT "DSCR1, overexpressed in Down syndrome, is an inhibitor of
RT calcineurin-mediated signaling pathways.";
RL Hum. Mol. Genet. 9:1681-1690 (2000).
RN [6]
RP FUNCTION: Inhibits calcineurin-dependent transcriptional responses
RN by binding to the catalytic domain of calcineurin A. Could play a
RN role during central nervous system development.
RN [7]
RP ALTERNATIVE PRODUCTS: 4 ISOFORMS; 1 (SHOWN HERE), 2, 3 AND 4; ARE
RN PRODUCED BY ALTERNATIVE SPLICING.
RN [8]
RP TISSUE SPECIFICITY: Highly expressed heart, brain and skeletal
RN muscle. Also expressed in all other tissues.
RN [9]
RP INDUCTION: By calcium.
RN [10]
RP SIMILARITY: BELONGS TO THE DSCR1 FAMILY.
RN [11]
RP This SWISS-PROT entry is copyright. It is produced through a collaboration
RN between the Swiss Institute of Bioinformatics and the EMBL outstation -
RN the European Bioinformatics Institute. There are no restrictions on its
RN use by non-profit institutions as long as its content is in no way
RN modified and this statement is not removed. Usage by and for commercial
RN entities requires a license agreement (see http://www.ebi.ac.uk/announce/
RN or send an email to license@ebi.ac.uk).

```

```

CC EMBL; U28833; AAB81557.1; -.
DR EMBL; U85265; AAB84370.1; -.
DR EMBL; U85266; AAB84371.2; -.
DR EMBL; U85267; AAB84372.1; -.
DR EMBL; BC002864; AAB02864.1; -.
DR EMBL; U53821; AAF21218.1; -.
DR GeneW; HGNC:3040; DSCR1.
KM MIM; 602917; -.
AL Alternative splicing.
FT DOMAIN 166 174 POLY-GLU.
FT VARSPLIC 1 28 MEEVDLDDPSATACHLDPVFDGLCR -> MHPFRNPNYS
FT VARSPLIC 1 29 FSLIACVANDISESET (IN ISOFORM 2).
FT VARSPLIC 1 29 MEEVDLDDPSATACHLDPVFDGLCR -> MYV (IN
FT VARSPLIC 1 80 ISOFORM 3).
FT VARSPLIC 1 80 MISSING (IN ISOFORM 4).
FT CONFLICT 159 159 H -> Q (IN REF. 4).
SQ SEQUENCE 197 AA; 22637 MW; 1BD426BB8167BEC CRC64;

Query Match 59.8%; Score 622; DB 1; Length 197;
Best Local Similarity 63.6%; Pred. No. 3e-47;
Matches 119; Conservative 26; Mismatches 34; Indels 8; Gaps 2;

QY 11 STLVAQVDVDEFTNOQVKEKEGELFRTYDECVTQQLFKSPRRVAINSHPSARARIE 70
DB 11 SATYCHLDPRVFDGLCRANFESLFRTYDKDITFYQKSFRRVAINSNPSAADARLQ 70
QY 71 LHETOPRGKLTLYPAQVOTPETDGDKLHAPPOPAKQFLSPSPSPVGMWPISDATPV 130
DB 71 LHKTEFLKEMKLYPAQTLHIGSS----HLAPRNDKQFLSPSPSPVGMWQVDAITPV 126
QY 131 INVDLYIAISKLGPEKXELHATDTTPSVVHVCSDEQENEEEMERKPKXIQT 186
DB 127 INVDLYIAISKLGPEKXELHATDTTPSVVHVCSDEQENEEEMERKPKXIQT 186
QY 187 RRRGLPP 193
DB 187 RREYTP 193

RESULT 7
CCPI_MOUSE
ID CCPI_MOUSE STANDARD; PRT; 198 AA.
AC Q9UHG6; Q9UKX3; Q9UKX2; Q9UK51; Q9UK50;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Calcipressin 1 (Down syndrome critical region protein 1 homolog)
DE (Myocyte-enriched calcineurin interacting protein 1) (MCIP1).
GN DSCR1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM A).
RX STRAIN=BA1B/C; TISSUE=Brain;
RN MEDLINE=20534792; PubMed=11080588;
RA Striopolis P., Petrini M., Lenzi L., Carinci P., Zamotti M.;
RT "The murine DSCR1-like (Down syndrome candidate region 1) gene family:
RN conserved synteny with the human orthologous genes.";
RL Gene 257:223-232 (2000).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS A AND B).
RX MEDLINE=20187590; PubMed=10722714;
RA Rothermel B., Vega R.B., Yang J., Wu H., Bassel-Duby R.,
RA Williams R.S.;
RT "A protein encoded within the Down syndrome critical region is
RN enriched in striated muscles and inhibits calcineurin signaling.";
RL J. Biol. Chem. 275:8719-8725 (2000).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM A), AND FUNCTION.
RX TISSUE=fetal brain;

```

Qy	11	STLVACVVDVEVFNTQNEVKKEFGLFRTYDECVTFOLFKSFRRVRNFSHPKSAARIE	70
Dd	11	SATIACHLDRPVFDGCLCRAKFESLFRTYDKOTFYFKSFEVRNFSNPISAADARLR	70
Qy	71	LHETOPRGGKLKYFAQVOTPTDGDKLHLAPPQAKQLISPSPPPVGWKPISDATPV	130
Dd	71	LHKTEFGKMKLYFAQTLLHGSS----HLAPPNDPKQLISPAPSPVPVGWKQVEDATPV	126
Qy	131	LNVDLLYAVAKLGEGKEYELHAGCTESTPVSVVHVCDSDMEEEDPK-----TSPKPKIIQ	185
Dd	127	INVDLLYAISKLGEGKEYELHAATDTTPSVVVHVCSDOENEEBEMRMKRPKPKIIQ	186
Qy	186	TRRGFLPP	193
Dd	187	TRREYTP	194

RESULT 8

ID	NLA_DROME	STANDARD;	PRT;	292 AA.
AC	OQXZL8; Q9V391;			
DT	30-MAY-2000 (Rel. 39, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Nebula protein.			
GN	NLA OR CG6072.			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;			
OC	Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;			
OC	Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.			
NCBI	TaxID=7227;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RP	McCormick A.V., Goldberg M.L.;			
RT	"Gene required for elongation of meiosis I spindle in Drosophila females."			
RL	Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.			
RP	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Berkely;			
RX	MEDLINE=20196006; PubMed=10731132;			
RA	Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,			
RA	Ananidis P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,			
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,			
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,			
RA	Stratton R.C., Rogers Y.-H.C., Blasej R.G., Champe M., Pfeiffer B.D.,			
RA	Abel J.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,			
RA	Wan K.J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,			
RA	Ballew R.M., Basu A., Bakendale J., Bayraktaroglu L., Beasley E.M.,			
RA	Beeson K.Y., Benos P.V., Bereman B.P., Bhandari D., Bolshakov S.,			
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotier P.,			
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,			
RA	Cherry J.M., Crawley S., Dahlke C., Davenport L.B., Davies P.,			
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,			
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,			
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,			
RA	Foeller C., Gabrielson A.B., Garg N.S., Gelbart W.M., Glasser K.,			
RA	Glocke A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,			
RA	Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,			
RA	Hootin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,			
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Kerchum K.A.,			
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai X.,			
RA	Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,			
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,			
RA	Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,			
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,			
RA	Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleby J.M.,			
RA	Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,			
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,			
RA	Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,			
RA	Spirer E., Spradling A.C., Stapleton M., Strong R., Sun E.,			
RA	Svirskas R., Testcor C., Turner R., Venter E., Wang A.H., Wang X.,			

Wang Z.-Y., Massarman D.A., Weinrock G.M., Weissenbach J., Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., Zheng X.H., Zhong F.N., Zhou W., Zhou X., Zhu S., Zhu X., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 "The genome sequence of *Drosophila melanogaster*.";
 Science 287:2185-2195(2000).
 RL -1- FUNCTION: REQUIRED FOR THE DSCR1 FAMILY.
 CC -1- SIMILARITY: BELONGS TO THE DSCR1 FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF147700; AAD33987.1; -
 CC EMBL: AE003712; AAF55285.1; -
 CC FlyBase: FBgn0026629; nla.
 DR PLYBASE: 292 AA; 31423 MW; 64F1BBF5FEAC6F9 CRC64;
 SQ SEQUENCE

Query Match 38.8%; Score 404; DB 1; Length 292;
 Best Local Similarity 43.2%; Pred. No. 4,2e-28;
 Matches 83; Conservative 35; Mismatches 62; Indels 12; Gaps 4;

4 PSMDCD-----VSTLVACVDEVFNTQVEKEKEGFLFTYDECVTQFLQKRRRLNF 58
 DB PEVDADSPDULFTSLITVNIHSEVFANPELKLAMELFTFESATFQWLRFRRLRVVY 163
 QY 59 SHPKSAARARIELHETQFGKK-LKLYPAQVQTPETDQKHLAPPAKQFLISPSPSP 117
 DB 164 DNALIAAANRILKHQYEFNKTKVTTCYFAQPTVPSN---KNLGPAPVYKQFLISPSP 220
 QY 118 PVGMKPISATVLYVDLLYAAKLGPGKXYLHAGTSTPSVVAHVCDMEBEDPRT 177
 DB 221 PAGWEPREGELVNHDLAALSLTPGSHSLHPOSEDOPAIYHTA---MLAETGPGL 277
 QY 178 SPKPIOTRRP 189
 DB 278 QVKAPIVOTKCP 289

RESULT 9
 CCPL CAEEL STANDARD: PRT; 207 AA.
 AC P53806; Q9U6V5; (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 41, Last annotation update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Calcipressin-like protein (Down Syndrome candidate region 1-like protein)
 DE RCN-1 OR DSCR1 OR PS4E7.7.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodidae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Bentley D.;
 RL Submitted (JUN-1994) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=20223370; PubMed=10756093;
 RA Strippoli P., Lenzi L., Petrini M., Zannotti M.;
 RT "A new gene family including DSCR1 (Down Syndrome candidate region 1) and ZAI-4: characterization from yeast to human and identification of DSCR1-like 2, a novel human member (DSCR1L2).";
 RL Genomics 64:252-263(2000).
 CC -1- FUNCTION: Inhibits calcineurin-dependent transcriptional responses

by binding to the catalytic domain of calcineurin A (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE DSCR1 FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U00067; AK20076.1; -
 CC EMBL: AF176115; AAF01683.1; -
 CC WormPep: F54E7.7; CE01318.
 DR CONFLICT 205 207
 FT CEQ -> SLOSSFYLCTPSSPIFFSPFCFNPQSLFCR
 FT G (IN REF. 1).
 SQ SEQUENCE 207 AA; 23030 MW; 0154B308AB05B79 CRC64;

Query Match 31.0%; Score 322.5; DB 1; Length 207;
 Best Local Similarity 38.9%; Pred. No. 3,4e-21;
 Matches 68; Conservative 36; Mismatches 64; Indels 7; Gaps 4;

11 STLVAQVDEVFNTQVEKEKEGFLFTYDECVTQFLKRRRLVINSHPKSAARARIE 70
 DB 30 NAIIVQPEDEVFDNKQKANSFLTQIEKDINHDFLRSRRRVVIFSSPENATPAKLI 89
 QY 71 LHETQFGKKLKLVAQVQTPETDQKHLAPPAKQFLISPSPPVGMKPISATV 130
 DB 90 VQGFPGKHELKAFPA--QRIYMSANSQMLSPPEKQFLISPSPSPVGMKPISATV 147
 QY 131 L-NYDLVAVAKLGPGKXYLHAGTSTPSVVAHVCDMEBEDPRTSPK 181
 DB 148 VCNFPLMARLASFAIDEXEYVNGDELTPAIVHPCETPIDVPAIEMPT-PRP 201

RESULT 10
 YAGF_SCHPO STANDARD: PRT; 163 AA.
 ID YAGF_SCHPO
 AC 009751;
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical protein C13G6.15c in chromosome I.
 GN SPAC13G6.15C OR SPAC24B11.04C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetes; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A., Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S., Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M., Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A., Gentles S., Goble A., Hamlin N., Harris D., Hladago J., Hodgson G., Holtroyd S., Hornsbly T., Howarth S., Huckle B.D., Hunt S., Jagsels K., James K., Jones L., Jones M., Leather S., McDonald D., McLean J., Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell J., Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E., Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S., Skelton J., Simmonds M., Squares R., Squares S., Stevens K., Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S., Woodward J., Voiclaert G., Aert R., Robben J., Grymopoulos B., Welter G., Vansireels E., Rieger M., Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Fritz C., Holzer E., Moesl D., Hilbert H., Borzsum K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M., Eger P., Zimmermann W., Medler H., Wandut R., Purnelle B., Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaune V., Mottier S., Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,

```

DR Genew; HGNC:16262; YAP1.
DR MIM; 606608; -.
DR InterPro; IPR005153; MbTH.
DR InterPro; IPR002349; WW.
DR InterPro; IPR001202; WW_Rep5_WWP.
DR Pfam; PF00397; WW; 1.
DR Pfam; PF03621; MbTH; 1.
DR PRINTS; PR00403; WWDOMAIN.
DR SMART; SM00456; WW; 1.
DR PROSITE; PS01159; WW_DOMAIN_1; 1.
DR PROSITE; PS50020; WW_DOMAIN_2; 1.
DR KEGG; Phosphorylation.
FT DOMAIN 171 204 WW.
SQ SEQUENCE 454 AA; 48755 MW; 87CB840D3393BFC0 CRC64;

Query Match 9.7%; Score 101; DB 1; Length 454;
Best Local Similarity 32.1%; Pred.No. 0.16; Matches 34; Conservative 10; Mismatches 30; Indels 32; Gaps
Qy 102 PPQAKQFLISPPSPVPGKPISDATPVLYNDLLYAVAKLGGKYEHLHAGTESTPSV- 160
Db 8 PPQAPQGGQSPQPPQGGPPS-----GPGQ--PAPAATQAPQAP 48
Qy 161 -----VVHV-CDSMDEEE-----DPKTSPPKIIQTRRGPPLPS 194
Db 49 PAGHQIVHVRGDSSETLEALFNAMVNPKTANVPQTVPMLRLKLPS 94

RESULT 12
CAP_YEAST
ID CAP_YEAST STANDARD; PRT; 526 AA.
AC P17555;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Adenylyl cyclase-associated protein (CAP).
GN SVR2 OR CAP1 OR YNL138W OR N1210 OR N1838.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OC NCBI_TaxID:4932;
RN 1[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90335281; PubMed=2184942;
RA Field J., Vojtek A., Ballester R., Bolger G., Colicelli J.,
RA Ferguson K., Gerst J., Kataoka T., Michaels T., Powers S., Riggs M.,
RA Rodgers L., Wieland L., Wheland B., Wiger M.;
RT "Cloning and characterization of CAP, the S. cerevisiae gene encoding
RL the 70 kd adenylyl cyclase-associated protein.";
RL Cell 61:319-327(1990).
RN 2[2]
RP SEQUENCE FROM N.A.
RX STRAIN=S288c;
RX MEDLINE=902335282; PubMed=2158860;
RA Fedor-Chaikin M., Deschenes R.J., Broach J.R.;
RT "SVR2, a gene required for RAS activation of adenylyl cyclase in
RL yeast.";
RL Cell 61:329-340(1990).
RN 3[3]
RP SEQUENCE FROM N.A.
RX STRAIN=S288c;
RX MEDLINE=961109332; PubMed=8619318;
RA Mallet L., Bussereau F., Jacquet M.;
RT "A 43.5 kb segment of yeast chromosome XIV, which contains MFA2,
RT MEF2, CAP/SRV2, NAM9, FKBP1/RBP1, MOM22 and CRT1, predicts an
RT adenosine deaminase gene and 14 new open reading frames.";
RL Yeast 11:1195-1209(1995).
CC -1- FUNCTION: THE N-TERMINAL DOMAIN BINDS TO ADENYLYL CYCLASE, THEREBY
CC ENABLING ADENYLYL CYCLASE TO BE ACTIVATED BY UPSTREAM REGULATORY
CC SIGNALS, SUCH AS RAS. THE C-TERMINAL DOMAIN IS REQUIRED FOR NORMAL
CC CELLULAR MORPHOLOGY AND GROWTH CONTROL.
CC -1- SUBCELLULAR LOCATION: CELL MEMBRANE.
CC -1- SIMILARITY: BELONGS TO THE CAP FAMILY.

```

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; 246843; CAA6687.1; -
 DR EMBL; M58284; AAB3569.1; -
 DR EMBL; M32653; AAB35094.1; -
 DR EMBL; 271414; CAA96020.1; -
 DR PIR; A34896; A34896.
 DR SGD; S0005082; SRV2.
 DR InterPro; IPR001837; CAP.
 DR Pfam; PF01213; CAP.1.
 DR PROSITE; PS01088; CAP.1; 1.
 DR PROSITE; PS01089; CAP_2; 1.
 KM Membrane.
 FT DOMAIN 262 300 ALA/PRO/SER-RICH.
 FT 277 282 POLY-PRO.
 SQ SEQUENCE 526 AA; 57521 MW; 0EB4D41205E2D464 CRC64;

Query Match 8.6%; Score 90; DB 1; Length 526;
 Best Local Similarity 24.9%; Pred. No. 1.7;
 Matches 52; Conservative 26; Mismatches 97; Indels 34; Gaps 10;

QY 10 VSTLVACVVD---VEVFTNQEVKEKEFGLFRYDECVTQFLFKSFRARRINPSHKSA 65
 Db 182 VTPVSMVVDLFDKDAQFVWYKRLIKE---YRESDP---NAVEWKKFLASDINKAYI 232
 QY 66 RARIELHETOPRGKLLKLYFAQVOTPETDGDKLHLAPPOPAKOFLL---SPSSPPVGM 121
 Db 233 K---EHTTGVSKKDKCMFPADMAOSTKNTGA-TSSPSASATAAPAPPPAPPAV 288
 QY 122 KPISDATPVLYNDL---LYAV-AKLGPGE-----KTELHAGTSTSVVHVCDSDM 169
 Db 289 FEISNDTPATSSDANKGIGAVFAELNCGENITKGLKVDKSOOTHKNDELROSSTVST 348
 QY 170 EEEEDPKTSP-KPKIOTRRPGLPPSVSN 197
 Db 349 GSKSGPPRPKPKSTLTKPKPKKELVGN 377

RESULT 13

RCN1_YEAST STANDARD; PRT; 211 AA.
 ID RCN1_YEAST
 AC P36054;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Calcipressin-1-like protein (Regulator of calcineurin) (Down Syndrome
 GN candidate region 1-like protein).
 GN RCN1 OR DSCR1L OR YK159C OR YK1613.
 OS Saccharomyces cerevisiae (Baker's Yeast).
 OS Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 NCBI_TaxID=4932;

KN SEQUENCE FROM N.A.
 RC STRAIN=6288C;
 RA MEDLINE=94378720; PubMed=8091859;
 RA Vandenbol M., Bolle P.-A., Dion C., Portetelle D., Hliger F.;
 RT "DNA sequencing of a 362 kb fragment located between the FAS1 and
 RT LAP1 loci of chromosome XI of Saccharomyces cerevisiae.";
 RL Yeast 10:535-540(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CG1945;
 RA MEDLINE=20221370; PubMed=10756093;
 RA Stripoli P., Lenzi L., Petrucci M., Carinci P., Zannotti M.;
 RT "A new gene family including DSCR1 (Down syndrome candidate region 1)

RT and ZAK1-4: characterization from yeast to human and identification of
 RT DSCR1-like 2, a novel human member (DSCR1L2).";
 RL Genomics 64:252-263(2000).
 RN [3]
 RP CHARACTERIZATION.
 RX MEDLINE=20347037; PubMed=10887154;
 RA Kingsbury T.J., Cunningham K.W.;
 RT "A conserved family of calcineurin regulators.";
 RL Genes Dev. 14:1595-1604(2000).
 CC -1- FUNCTION: Inhibits calcineurin-dependent transcriptional responses
 CC by binding to the catalytic domain of calcineurin.
 CC -1- SIMILARITY: BELONGS TO THE DSCR1 FAMILY.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; 226877; CAA81495.1; -
 DR EMBL; 228159; CAA82001.1; -
 DR EMBL; AF174139; AAF01680.1; -
 DR PIR; S37792; S37792.
 DR PIR; S44569; S44569.
 DR SGD; S0001642; RCN1.
 SQ SEQUENCE 211 AA; 24135 MW; 668EA0C10DAE39A CRC64;

Query Match 8.4%; Score 87.5; DB 1; Length 211;
 Best Local Similarity 25.2%; Pred. No. 0.92; Mismatches 38; Indels 31; Gaps 6;
 Matches 31; Conservative 23;

QY 8 CDVSTLVACVVD---VEVFTNQEVKEKEFGLFRYDECVTQFLFKSFRARRI----- 56
 Db 16 CD-----IVDNDVVERIQWLSKNILRKFG--INENEPLOQLILKRFKILLICPSH 65
 QY 57 -NFSPPKAAARIRLHTOPRGKLLKLYFAQVOTPETDGDKLHLAPPOPAKOFLLSPS 115
 Db 66 DISQHVMDASRA-LEMENFNS-----YSLDGGQNLTKOYLKVESEKMFLLSPPA 116
 QY 116 SPP 118
 Db 117 SPP 119

RESULT 14

MAP2_HUMAN STANDARD; PRT; 1827 AA.
 ID MAP2_HUMAN
 AC P11137; Q99976; Q99975.
 DT 01-JUN-1989 (Rel. 11, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Microtubule-associated protein 2 (MAP2) [Contains: MAP2C].
 GN MAP2.
 OS Homo sapiens (Human).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Price R.;
 RL Submitted (SEP-1993) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
 RC TISSUE=Brain;
 RX MEDLINE=94124038; PubMed=8294038;
 RA Albala J.S., Kalcheva N., Shafit-Zagardo B.;
 RT "Characterization of the transcripts encoding two isoforms of human
 RT microtubule-associated protein-2 (MAP-2).";
 RL Gene 136:377-378(1993).
 RN [3]
 RP SEQUENCE OF 493-1562 FROM N.A.


```

RX MEDLINE=88274407; PubMed=2455776;
RA Kosik K.S., Orcicchio L.D., Bakalis S., Duffy L., Neve R.L.;
RT "Partial sequence of MAP2 in the region of a shared epitope with
RT Alzheimer neurofibrillary tangles.";
RL J. Neurochem. 51:587-598(1988).
CC -!- FUNCTION: THE EXACT FUNCTION OF MAP2 IS UNKNOWN BUT MAPS MAY
CC STABILIZE THE MICROTUBULES AGAINST DEPOLYMERIZATION. THEY ALSO
CC SEEM TO HAVE A STIFFENING EFFECT ON MICROTUBULES.
CC -!- ALTERNATIVE PRODUCTS: VARIOUS FORMS OF MAP2 ARE PRODUCED BY
CC ALTERNATIVE SPLICING OF THE SAME GENE. MAP2C, THE LOW MOLECULAR
CC FORM OF MAP2, LACKS THE CENTRAL DOMAIN OF MAP2A/B.
CC -!- SIMILARITY: CONTAINS 3 TAU/MAP REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U01828; AAA03354.1; -
CC EMBL; U89330; AAB48098.1; -
CC EMBL; U89329; AAB48097.1; -
CC EMBL; M25668; AAA59552.1; -
CC PIR; PL0024; QRHUMT.
CC Genew; HGNC:6839; MAP2.
CC MIM; 157130; -
CC InterPro; IPR001084; Tubulin Tau.
CC Pfam; PF00418; tubulin-binding; 3.
CC PROSITE; PS000229; TAU_MAP; 2.
KW Microtubules; Repeat; Alternative splicing; Calmodulin-binding.
FT DOMAIN 1447 1467 CALMODULIN-BINDING (POTENTIAL).
FT REPEAT 1661 1691 TAU/MAP MOTIF.
FT REPEAT 1692 1722 TAU/MAP MOTIF.
FT REPEAT 1723 1754 TAU/MAP MOTIF.
FT VARSPLIC 152 1507 MISSING (IN ISOFORM MAP2C).
FT CONFLICT 9 9 A -> G (IN REF. 2).
FT CONFLICT 37 37 R -> A (IN REF. 2).
FT CONFLICT 108 108 A -> G (IN REF. 2).
FT CONFLICT 152 155 MISSING (IN REF. 2).
FT CONFLICT 187 187 S -> K (IN REF. 2).
FT CONFLICT 1655 1655 A -> GL (IN REF. 2).
FT CONFLICT 1736 1736 A -> A (IN REF. 2).
SQ SEQUENCE 1827 AA; 199610 MW; BAC36D0030F5F455 CRC64;

Query Match 8.1%; Score 84.5; DB 1; Length 1827;
Best Local Similarity 35.9%; Pred. No. 23;
Matches 28; Conservative 6; Mismatches 31; Indels 13; Gaps 3;

QY 101 APPQAKQFLISPPSPVGVGWKPISDATPVLYNLYAVAKLGPGEKVELHAGTSTPSV 160
Db 122 ALPLAAEETANLPPSPPP---SPASEQTVTVEDLLTA-----SGMEFHDQQLTPST 171

QY 161 VHVCDSDMEEEEDPKTS 178
Db 172 AE---PSDQKEKESKQS 186

RESULT 15
FH14_MOUSE
ID FH14_MOUSE STANDARD; PRT; 1206 AA.
AC Q05859;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Formin 1 isoform IV (limb deformity protein).
GN FMN OR LD.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathii; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RX TISSUE=Embryo;
RC MEDLINE=92112033; PubMed=1339380;
RT Grubby-Jackson L., Kuo A., Leder P.;
RA "A variant limb deformity transcript expressed in the embryonic mouse
RT limb defines a novel formin.";
RL Genes Dev. 6:29-37(1992).
CC -!- FUNCTION: IS IMPORTANT IN THE MORPHOGENESIS OF LIMB AND MAY HAVE A
CC FUNCTION IN DIFFERENTIATED CELLS OR BE INVOLVED IN MAINTAINING
CC SPECIFIC DIFFERENTIATED STATES.
CC -!- ALTERNATIVE PRODUCTS: AT LEAST 5 ISOFORMS: IA (AC Q05860), IB (AC
CC Q05860), II (AC Q05860), III (AC Q05860) AND IV (SHOWN HERE); ARE
CC PRODUCED BY ALTERNATIVE SPLICING. A VARIATION IN SPLICING IS SEEN
CC AMONG DIFFERENT TISSUES AND DIFFERENT SIZE TRANSCRIPTS EXIST
CC WITHIN ANY ONE TISSUE.
CC -!- TISSUE SPECIFICITY: IT IS FOUND THROUGHOUT THE EMBRYO BUT
CC HAS A FUNCTIONAL ROLE ONLY IN THE KIDNEY AND LIMB.
CC -!- DEVELOPMENTAL STAGE: THIS IS THE ISOFORM FOUND IN THE APICAL
CC ECTODERMAL RIDGE AND THE MESENCHYMAL COMPARTMENT OF THE DEVELOPING
CC LIMB BUD.
CC -!- PTM: PHOSPHORYLATED ON SERINE AND POSSIBLY THREONINE RESIDUES.
CC -!- SIMILARITY: CONTAINS 1 FORMIN HOMOLGY 1 (FH1) DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 FORMIN HOMOLGY 2 (FH2) DOMAIN.
CC -!- SIMILARITY: BELONGS TO THE FORMIN HOMOLGY FAMILY. CAPPUCCINO
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X62379; CAA44244.1; -
CC PIR; S24407; S24407.
CC MGD; MGI:101815; Fmn.
CC InterPro; IPR003104; FH2.
CC InterPro; IPR001265; Formin.
CC Pfam; PF02181; FH2; 1.
CC PRINTS; PR00828; FORMIN.
CC SMART; SM00498; FH2; 1.
KW Nuclear protein; Developmental protein; Alternative splicing;
KW Phosphorylation; Coiled coil.
FT DOMAIN 418 443 COILED COIL (POTENTIAL).
FT DOMAIN 457 566 COILED COIL (POTENTIAL).
FT DOMAIN 644 744 FH1 (PRO-RICH).
FT DOMAIN 759 1164 FH2.
FT DOMAIN 1043 1116 COILED COIL (POTENTIAL).
FT DOMAIN 635 638 POLY-SER.
FT DOMAIN 751 755 POLY-SER.
SQ SEQUENCE 1206 AA; 133464 MW; 4DFB38CB52BD8EE7 CRC64;

Query Match 8.0%; Score 83; DB 1; Length 1206;
Best Local Similarity 23.4%; Pred. No. 19;
Matches 46; Conservative 18; Mismatches 87; Indels 46; Gaps 8;

QY 26 QEVKEFEGFLRTYDECVTFFQ---LFKSPRV-----RINFSPKSAARARIELHETQFR 77
Db 547 QQLEKRECEEMRDVCISTDDDCSPKAFNVCIQDRETFLKPCDAESK--ATRSQIV 604

QY 78 GKKLKLYFAQVQTPETDGDGLHL-----AP---PPAKQFLISPPSS 116
Db 605 PKKLITSLTQL-SPKSDKDIHAPFOTREGTSSSSQOKISPPAPPTPLPPPLIPPPP 663
QY 117 PPVGWKPISDATPVLYNLYAVAKLGPGEKVELHAGTSTPSVHVCDSDMEEEEDPK 176
Db 664 LPPGLGLPLPPAPPI-----PPVCPVSP-----PPPPPPPPPTVPESDPPPPPPP 709
QY 177 TSPKPKIITRRPGPLPP 193
Db 710 PPPLPNVLALPNSGGPP 726

```

Sun Dec 15 08:39:01 2002

us-09-782-953-9.rsp

Page 10

Search completed: December 11, 2002, 11:39:43
Job time : 9.3086 secs

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	322.5	31.0	239	2	T34305	hypothetical prote
2	133.5	12.8	315	2	T49641	nebula related pro
3	107.5	10.3	163	2	S62444	conserved hypothet
4	101	9.7	454	2	A55954	yes-associated pro
5	93	8.9	1018	2	JC4211	neural adhesion pr
6	92	8.8	277	2	T37629	hypothetical prote
7	90	8.6	526	2	A34896	adenylate cyclase-
8	87.5	8.4	211	2	S37792	hypothetical prote
9	87.5	8.4	13055	2	T15850	hypothetical prote
10	87	8.4	1018	2	T19693	hypothetical prote
11	86	8.3	1006	2	T4731	atrophin-1 related
12	85.5	8.2	7576	2	T17428	FK506 polypeptide s
13	84.5	8.1	212	2	T48580	hypothetical prote
14	84.5	8.1	1222	2	B95933	hypothetical prote
15	84.5	8.1	3938	2	T42761	Bassoon protein -
16	84	8.1	671	2	C96534	probable Poly-A Bi
17	84	8.1	1557	2	T13160	protein CNK - frui
18	83	8.0	160	2	C72706	hypothetical prote
19	83	8.0	643	2	T23453	hypothetical prote
20	83	8.0	746	2	S74546	glycogen operon pr
21	83	8.0	1206	2	S24407	formin isoform IV
22	83	8.0	1232	2	S40766	hypothetical prote
23	83	8.0	1468	2	S11515	formin - mouse
24	82.5	7.9	656	2	G86482	protein FSJ5.9 (im
25	82	7.9	1611	2	T38236	hypothetical prote
26	82	7.9	2464	1	QRWSP1	microtubule-associ
27	82	7.9	2783	1	A41948	alpha-fetoprotein
28	81	7.8	559	2	A57474	extracellular matr
29	81	7.8	913	2	A73203	cellobiose-phospho

Db 913 SSVRSGRYIITWDHVVALSNESTVTGKYLPRPDGQHDGKLYSTHKHSIEVP 965

RESULT 6

T37629
hypothetical protein SPAC13f5.04c - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T37629
R:Brown, D.; Churcher, C.M.; Barrrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, September 1997
A:Reference number: Z21733
A:Accession: T37629
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-277 <BRO>
A:Cross-references: EMBL:Z99091; PIDN:CAB11767.1; GSPDB:GN00066; SPDB:SPAC13f5.04c
A:Experimental source: strain 972h-; cosmid c13f5
C:Genetics:
A:Gene: SPDB:SPAC13f5.04c
A:Map position: 1
A:Introns: 20/3

Query Match 8.8%; Score 92; DB 2; Length 277;
Best Local Similarity 22.0%; Pred. No. 0.66; 76; Indels 60; Gaps 7;
Matches 45; Conservative 24; Mismatches 24

QY 42 CVTFQFQKSRVRINF-----SHIPKSAAR-----ARIELHETQFRGK 79

Db 2 CFFFLVSSPAILRISLTSIKFKCKLRALQVANPQRKAKTPSNHATELQSSSTNST 61

QY 80 KKLKLYFAQVOTPDGDKLHLAPP--OPAKQFIISPSSPPVGMWK-----ISDA 127

Db 62 TLPTQEAAVETNASHETSFALETPSPAASLSISPTKSAASVSEPNVEADVKSLSSTPA 121

QY 128 TPVLNVDLLYAVAKLGPEKYE-----LHAGTESPSV-----VVHVCDSMEE 171

Db 122 APQLN-----SPSHSYETTFPTSTTSITENLPTIDPTSTRSSHIQSLSPESQ 171

QY 172 EEDPKTSPKPKIIOTRRPGLPPSVS 196

Db 172 TSDGHRPPSPSTSTTSIDPSVA 196

RESULT 7

A34896
adenylate cyclase-associated protein, 70K - yeast (Saccharomyces cerevisiae)
N:Alternate names: adenylate cyclase-binding protein CAP; protein JTA526; protein N1210;
C:Species: Saccharomyces cerevisiae
C>Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 23-Mar-2001
C:Accession: A34896; A34897; S55145; S59250; S63083
R:Field, J.; Vojtek, A.; Ballester, R.; Bolger, G.; Collicelli, J.; Ferguson, K.; Gerst, C.
Cell 61, 319-327, 1990
A:Title: Cloning and characterization of CAP, the S. cerevisiae gene encoding the 70 kD
adenylate cyclase-associated protein, 70K - yeast (Saccharomyces cerevisiae)
A:Reference number: A34896; MUID:90235281; PMID:2184942
A:Accession: A34896
A:Molecule type: mRNA
A:Residues: 1-526 <PIE>
A:Cross-references: EMBL:M58284; NID:g171156; PIDN:AAA63569.1; PID:g171157
R:Fedor-Chaikin, M.; Deschenes, R.J.; Broach, J.R.
Cell 61, 329-340, 1990
A:Title: SRV2, a gene required for RAS activation of adenylate cyclase in yeast.
A:Reference number: A34897; MUID:90235282; PMID:2158860
A:Accession: A34897
A:Molecule type: DNA
A:Residues: 1-526 <PED>
A:Cross-references: EMBL:M32663; NID:g172708; PIDN:AAA35094.1; PID:g172709
R:Mallet, L.; Bussereau, F.; Jacquet, M.
submitted to the EMBL Data Library, November 1994
A:Description: A 43.5 kb fragment of the chromosome XIV.
A:Reference number: S55136
A:Accession: S55145

A:Molecule type: DNA

A:Residues: 1-526 <MAL>

A:Cross-references: EMBL:Z46843; NID:g861113; PIDN:CAA86887.1; PID:g954499
R:Mallet, L.; Bussereau, F.; Jacquet, M.
Yeast 11, 1195-1209, 1995

A:Title: A 43.5 kb segment of yeast chromosome XIV, which contains MFA2, MEP2, CAP/SRV2
submitted to the EMBL Data Library, November 1994
A:Reference number: S59241; MUID:96109932; PMID:8619318
A:Accession: S59250

A:Molecule type: DNA

A:Residues: 1-526 <MAV>

A:Cross-references: EMBL:Z46843; NID:g861113; PIDN:CAA86887.1; PID:g954499
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1994
R:Mallet, L.; Bussereau, F.; Jacquet, M.

A:Reference number: S63069

A:Accession: S63083

A:Molecule type: DNA

A:Residues: 1-526 <MAF>

A:Cross-references: EMBL:Z71414; NID:g1302083; PIDN:CAA96020.1; PID:g1302084; MIPS:YNL1
A:Experimental source: strain S288C
C:Genetics:

A:Gene: SGD:SRV2; CAP1

A:Cross-references: SGD:S0005082; MIPS:YNL138W

A:Map position: 14L

C:Superfamily: adenylyl cyclase-associated protein MCH1

Query Match

Best Local Similarity 8.6%; Score 90; DB 2; Length 526;

Matches 52; Conservative 26; Mismatches 97; Indels 34; Gaps 10;

QY 10 VSTLVACVVD----VEVFTNQEVKEKFEGLFTYDECVTQFLKFSRRVRINFHPKSA 65

Db 182 VDTPSVMVTDPKDAAQFWINRLKE-----YRESDP-----NAVEVWKFLASFDNLKAYI 232

QY 66 RARIELHETQFRGKLLKLYFAQVOTPDGDKLHLAPPQAKQFI-----SPSSPPVGM 121

Db 233 K---BYHTTGVSWKDKGMDFAADAMAQSTKNTGA--TSSPSPASATAAPAPPPPPAPPASV 288

QY 122 KPISDATPVNLYDL---LYAV-AKLGPCE-----KVELHAGTESPSVVVHVCDSM 169

Db 289 FEISNDTPATSSDANKGGIGAVFAELNOGENTKGLKKVDKSKQOQTHKNPELROSSTVSST 348

QY 170 EEEEDPKTSP-KPKIIOTRRPGLPPSVSN 197

Db 349 GSKSGEPPEPKPSTLTKRPPRKELVGN 377

RESULT 8

S37792

hypothetical protein YKL159c - yeast (Saccharomyces cerevisiae)

N:Alternate names: hypothetical protein YKL613

C:Species: Saccharomyces cerevisiae

C>Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 19-Apr-2002

C:Accession: S37792; S37989; S44569

R:Vandenbol, M.; Bolle, P.; Dion, C.; Portetelle, D.; Hilger, F.

submitted to the EMBL Data Library, September 1993

A:Description: DNA sequencing of a 36.2 kb fragment located between the FAS1 and LAP4 loci

A:Reference number: S37786

A:Accession: S37792

A:Molecule type: DNA

A:Residues: 1-211 <VAN>

A:Cross-references: EMBL:Z26877; NID:g407482; PIDN:CAA81495.1; PID:g407489

A:Experimental source: strain S288C

R:Vandenbol, M.; Bolle, P.A.; Dion, C.; Portetelle, D.; Hilger, F.

submitted to the Protein Sequence Database, March 1994

A:Reference number: S37976

A:Accession: S37989

A:Molecule type: DNA

A:Residues: 1-211 <VA2>

A:Cross-references: EMBL:Z28159; NID:g486276; PIDN:CAA82001.1; PID:g486277; MIPS:YKL159C

A:Experimental source: strain S288C

R:Vandenbol, M.; Bolle, P.A.; Dion, C.; Portetelle, D.; Hilger, F.

Yeast 10, 35-40, 1994
 A:Title: DNA sequencing of a 36.2 kb fragment located between the PAS1 and LAP4 loci of
 A:Reference number: S44563
 A:Accession: S44563

A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-211 <VA3>

A:Cross-references: EMBL:Z26877; NID:G407482; PIDN:CAA81495.1; PID:G407489

A:Experimental source: strain S288C
 A>Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1993

C:Genetics:

A:Gene: SGD:RCN1

A:Cross-references: SGD:S0001642

A:Map position: 11L

C:Superfamily: Saccharomyces cerevisiae hypothetical protein YKL159C

Query Match 8.4%; Score 87.5; DB 2; Length 211;
 Best Local Similarity 25.2%; Pred. No. 1.2; Mismatches 38; Indels 31; Gaps 6;
 Matches 31; Conservative 23;

QY 8 CDVSTLVACVD-----VEFTNQVEKKEFGLFRTYDECVTFFQKSPRRVRI----- 56

DB 16 CD-----IVDNDNERIGVWLSKILRFG---INENBPIQLILKKPKRIILICPSH 65

QY 57 -NFSHPKSAARARIELHETQFRGKULKYFAOVQTPETDGDKLHLAPQPAKQFLISPS 115

DB 66 DISQHVMDASRA-LEMENFNFS-----YSLQDQORNLTKQYLKVPSEKMFILISPA 116

QY 116 SPP 118
 DB 117 SPP 119

Result 9

hypothetical protein K07E12.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C:Accession: T16580

R:Fullon, L.

submitted to the EMBL Data Library, May 1994

A:Description: The sequence of C. elegans cosmid K07E12.

A:Reference number: Z18540

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Accession: T16580

A:Molecule type: DNA

A:Residues: 1-1305 <FLL>

A:Cross-references: EMBL:U00054; NID:G485140; PID:G485141; PIDN:AAA50715.1; CESP:K07E12.

A:Experimental source: strain Bristol N2

C:Genetics:

A:Gene: CESP:K07E12.1

A:Introns: 46/3; 85/2; 201/1; 278/2; 470/2; 817/1; 927/1; 960/3; 1265/1; 1322/1; 1478/3; 6014/3; 6159/3; 6665/2; 7266/3; 7895/3; 8726/3; 9803/3; 10937/3; 12234/2; 1229

Query Match 8.4%; Score 87.5; DB 2; Length 1305;
 Best Local Similarity 23.4%; Pred. No. 2.2e+02;
 Matches 54; Conservative 27; Mismatches 79; Indels 71; Gaps 11;

QY 1 MPAPSMDCVSTLVACVDVEFTN-----OEVEKKEGLR-tydecvt-fol 47

DB 447 MTFEIDCNVTR--ADVPVVLWHLKGRPLKSGSKTOHIMKKGCVLESTOFSCVAENEA 504

QY 48 FKSPRRVIRINFSHPKSAARARIELHETQFRGKULKYFAOVQTPETDGDKLHL-APPQ 105

DB 505 GKSTKILNIVTGPSAERIRIYQ-----DGDVTLQWEPQI 542

QY 106 AK-----OFLSPSPSPVGMK-----PISDAPVLAVD-----LLYAVAKLGP 145

DB 543 TNGPMAGVDVFTEDPSLPRDQWKHHIDDPNARTTTLRLNEKTPYTFVIYVGRNRLPG 602

QY 146 ---EKYLHACTESTPSVVVHVCSDMEEBD-----PKTSPPKII 184

DB 603 LPSAPFTATTWLAKPPVQLEPSEMTKEPSNDEMIIECGAGCVKPKII 653

Result 10

hypothetical protein C34B7.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T19693

R:Haris, B.

submitted to the EMBL Data Library, December 1996

A:Reference number: Z19165

A:Accession: T19693

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1018 <WIL>

A:Cross-references: EMBL:Z83220; PIDN:CA805700.1; GSPDB:GN00019; CESP:C34B7.1

A:Experimental source: clone C34B7

C:Genetics:

A:Gene: CESP:C34B7.1

A:Map position: 1

A:Introns: 78/1; 149/1; 177/1; 245/1; 271/1; 311/2; 361/2; 387/3; 415/2; 444/3; 486/2; 5

Query Match 8.4%; Score 87; DB 2; Length 1018;
 Best Local Similarity 24.0%; Pred. No. 9.8;
 Matches 46; Conservative 23; Mismatches 85; Indels 38; Gaps 8;

QY 1 MPAPSMDCV-----TLVACVDVEFTNQVEKKEFGLFRTYDECVTFFQKFS 50

DB 337 LPTPAPQNSVSKVCNABILITLLFTKVEKSVSPSLNPEPLESLFRNPANCFENIFYL 396

QY 51 F--RRVIRINFSHPKSAARARIELHETQFRGKULKYFAOV---QTPETDGDK-LHLAP 103

DB 397 IFNRKMGKSTKPTSKANSSPE-----GRKSPMKTQLDDQOTPSSSSKKNLKKP 449

QY 104 QPAKQFLISPPSPVGMKPISDATPVNLVLDLYAVAKLGPGEYELHACTESTPSVVH 163

DB 450 STDD-----TPSSSEPLSKNLDSTETAEED-----NTRRRRTLKGVALP--VCS 495

QY 164 VCDSEEBEDP 175

DB 496 FCKEDQKADDEP 507

Result 11

atrophin-1 related protein - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 11-Jan-2000

C:Accession: T42731

R:Khan, F.A.; Margolis, R.L.; Loev, S.L.; Sharp, A.H.; Li, S.H.; Ross, C.A.

submitted to the EMBL Data Library, December 1995

A:Description: CDNA sequence and expression of an atrophin-1 (DRPLA disease gene) related

A:Reference number: Z22250

A:Accession: T42731

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1006 <KIA>

A:Cross-references: EMBL:U44091; NID:G1297310; PID:G1209103; PIDN:AAA98970.1

A:Gene: ARP

Query Match 8.3%; Score 86; DB 2; Length 1006;
 Best Local Similarity 28.2%; Pred. No. 12;
 Matches 33; Conservative 13; Mismatches 43; Indels 28; Gaps 7;

QY 91 PENTGDKLHLP---POPAKQFL-----ISPPSPVGMKPI---SDATPVNLVLDLYA 138

DB 459 PTTGG--LHQVDSQSPFPHFVPGGPPPTTPSPSCPPTSPAGSSSSQPCSS----- 510

QY 139 VAKLGPGEYELHACTESTPSVVVHVCSDMEEBDPKTSPPKXIIQTRRPLGPSV 195

DB 511 -AAVSSGNAV---PGAPSCPLPAVQIKEXALDEABEPSPPPP---PRSPPEPTV 559

RESULT 12

T17428
FK506 polyketide synthase - Streptomyces sp. (strain MA6548)
C:Species: Streptomyces sp.
A:Variety: strain MA6548
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Nov-2000
C:Accession: T17428
R:Motamedi, H.; Shafiee, A.
Eur. J. Biochem. 256, 528-534, 1998
A:Title: The biosynthetic gene cluster for the macrolactone ring of the immunosuppressant FK506 polyketide synthase
A:Reference number: 218779; MUID:98451508; PMID:9780228
A:Accession: T17428
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-7576 <MOT>
A:Cross-references: EMBL:AF082100; NID:G3798623; PID:G3798624; PIDN:AAC68815.1
A:Experimental source: strain MA6548
C:Genetics:
A:Gene: fkbB
C:Function:

A:Description: involved in synthesis of the backbone of the immunosuppressant FK506 polyketide
C:Superfamily: 3-oxoacyl-[acyl-carrier-protein] synthase I homology; acetate-CoA ligase
C:Keywords: carrier protein
F:54-500/Domain: acetate-CoA ligase homology <ACL>
F:1095-1166/Domain: acyl carrier protein homology <ACP1>
F:1204-1599/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS1>
F:1682-1953/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT1>
F:2680-2751/Domain: acyl carrier protein homology <ACP2>
F:2804-3198/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS2>
F:3295-3569/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT2>
F:4320-4391/Domain: acyl carrier protein homology <ACP3>
F:4435-4830/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS3>
F:5903-5974/Domain: acyl carrier protein homology <ACP4>
F:6018-6412/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS4>
F:6513-6785/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT3>
F:7450-7521/Domain: acyl carrier protein homology <ACP5>

Query Match 8.2%; Score 85.5; DB 2; Length 7576;
Best Local Similarity 26.9%; Pred. No. 1.7e+02;
Matches 39; Conservative 17; Mismatches 60; Indels 29; Gaps 7;

QY 47 LFPSFRVRINFS-----HPKSAARARIELHETQRGKKLYFAQVQ--TPETDGDKLH 99

Db 3834 LFADPERRLAFWSGVRIHARAATQTVLLERAD--GDTIRILATDEHGAPVLDVGLT 3891

QY 100 LAPPQPAKQEL-----ISPPSPVGVGKPTSDATPVNLVLLYAVAKLGPGEKVELHAGT 154

Db 3892 VRAAEPTGTEALFVAMVSVFVPGWTHLAD---VPEGDVL-----EGDAGEGDVGE 3941

QY 155 ESTPSVVVHVCDSDMEEDPKTSP 179

Db 3942 SDPPVVVVLAV-----EPGDPDSSP 3961

RESULT 13

T48580
hypothetical protein T31B5.150 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C:Accession: T48580
R:Bevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Bancroft, submitted to the Protein Sequence Database, April 2000
A:Reference number: Z24490
A:Accession: T48580
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-212 <BEV>
A:Cross-references: EMBL:AL163491
A:Experimental source: cultured Columbia; BAC clone T31B5
C:Genetics:
A:Map position: 5
A:Introns: 32/1

A>Note: T31B5.150

Query Match 8.1%; Score 84.5; DB 2; Length 212;
Best Local Similarity 26.9%; Pred. No. 2.3;
Matches 35; Conservative 14; Mismatches 38; Indels 43; Gaps 7;

QY 61 PKSAARARIELHET-----QPRGKKLYFAQVQ--VOTPETD-----GD 96

Db 57 PKKAARVWLGTFTAEAAALAYDRAALKFKGTAKLNFFERVQGPPTTTTISHAPRGVSE 116

QY 97 KULHAPPQPAKQFLISPPSPVGVGKPTSDATPVNLVLLYAVAKLGPGEKVELHAGTE- 155

Db 117 SMNSPPRP-----GPPSTTTTSW-PMT-----YNQDILQYLAQLITSNNEVDLSVYTTST 164

QY 156 -----STPS 159

Db 165 LFSQFPSTPS 174

RESULT 14

B90593

hypothetical protein MYPV 6500 [imported] - Mycoplasma pulmonis (strain UAB CTIP)

C:Species: Mycoplasma pulmonis

C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 03-Aug-2001

C:Accession: B90593

R:Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, I., Nucleic Acids Res. 29, 2145-2153, 2001

A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulmonis

A:Reference number: A99512; MUID:21267165; PMID:11353084

A:Accession: B90593

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1222 <KUR>

A:Cross-references: GB:AL445566; PID:gl4090065; PIDN:CAC13823.1; GSPDB:GN00153

A:Experimental source: strain UAB CTIP

C:Genetics:

A:Gene: MYPV 6500

A:Genetic code: SGC3

Query Match

Best Local Similarity 8.1%; Score 84.5; DB 2; Length 1222;
Matches 41; Conservative 34; Mismatches 68; Indels 43; Gaps 8;

QY 9 DVSTLVACVDDVEFTNOEVKEKPEGL---PRTYDECVTQFQKFSRRVRINFSPKSA 65

Db 218 DVTSASAKIRVNAKLTNNINNKFEVINGFKY--VIADSLKAKFTLVK-----KDIA 269

QY 66 RARIELHETO-FRGKKLYFAQVQTPETDGDKLHAPPQPAKQFLISPPSPVGVGKPI 124

Db 270 IBEKRVVTEAFKSKTQEFASLK-----NYLDISLPEGLDPSFVS-----IETKND 317

QY 125 SDATPVNLVLLYAVAKLGPGEKVELHAGTSTPSVVHVCDSDME-----EEDDP 175

Db 318 DPESAILTYKLVK-----NLHSGSQANPSKNGVSVESKLKTFEIKRSQIQE 367

RESULT 15

T42761

Bassoon protein - rat

N:Alternate names: brain-specific synapse-associated protein

C:Species: Rattus norvegicus (Norway rat)

C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 21-Jul-2000

C:Accession: T42761

R:Dieck, S.; Sanmarti-Vila, L.; Langnaese, K.; Richter, K.; Kindler, S.; Soyke, A.; Wex, J. Cell Biol. 142, 499-509, 1998

A:Title: Bassoon, a novel zinc-finger CAG/Glutamine-repeat protein selectively localized

A:Reference number: Z22499; MUID:98345363; PMID:9679147

A:Accession: T42761

A:Status: preliminary; translated from GB/EMBL/DBJ

Query Match	29.1%	Score 84.5	DB 2	Length 3938	
Best Local Similarity	8.1%	Pred. No. 91			
Matches 34	Conservative 17	Mismatches 36	Indels 27	Gaps 8	

```

QY      66  AOVURPETHDAKYNLAPROPKOKLSPSPSPRYMIMEDIDAPRYNLYDLYAAVLYSG 145
DB      1441 ASREKRLSGGSG-EGGPRQPSRGVSYPTSGSP-----PLSPSP--SESPTSPSPKCP 1492
QY      146  EKKYLNATES-TPSVVUNHCDSSMEEDBPKT--SPKPKTIG-----TRRQPLP 192
DB      1493 ATAESITQRYLTYS-----SLT-----PRVSGTSPVNAQSTQYNNRSESTR 1534

```

Search completed: December 11, 2002, 11:38:40
Job time : 19.9562 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

QM protein - nucleic search, using frame_plus_p2n model

Run on: December 14, 2002, 18:14:56 ; Search time 1647.42 Seconds
(without alignments)
3480.130 Million cell updates/sec

Title: US-09-782-953-9

Perfect score: 1041

Sequence: 1 MFAPSMDCVSTLVACVVDV.....SPKPKIQRFRGLPPSVSN 197

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4103280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO spoof/US09782953/runat_11122002_114429_17340/app query.fasta_1.1173
-DB=GenEmbl -QFMT=fastap -SUFFIX=range -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=i -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pcio -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09782953 @CNC 1.1.2844 @runat_11122002_114429_17340 -NCPU=6 -ICPU=3
-NO_XLPXY -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_scs.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_ox.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*

29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rnd.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1041	100.0	594	6	AX365318 Sequence
2	1041	100.0	594	10	AF237791 Mus muscu
3	1041	100.0	659	10	AF237887 Mus muscu
4	1027	98.7	3227	10	AB061524 Mus muscu
5	1007	96.7	3184	6	AX329596 Sequence
6	1007	96.7	3184	6	AX333015 Sequence
7	1007	96.7	3184	6	AX365327 Sequence
8	1007	96.7	3184	6	AX420436 Sequence
9	1007	96.7	3184	9	HUMZAKI4
10	926	89.0	3295	10	AB061525 Mus muscu
11	910	87.4	1021	6	AX420427 Sequence
12	905	86.9	934	6	AX074351 Sequence
13	905	86.9	3159	6	AX420425 Sequence
14	905	86.9	3261	9	AK090990 Homo sapi
15	892	85.7	3240	9	AY034085 Homo sapi
16	892	85.7	3253	9	AY034086 Homo sapi
17	663.5	63.7	776	10	AF237888 Mus muscu
18	658.5	63.3	828	6	AX365330 Sequence
19	658.5	63.3	828	6	AX420437 Sequence
20	658.5	63.3	828	9	AF176116 Homo sapi
21	643	61.8	626	10	AB075973 Rattus no
22	641	61.6	2216	10	CGU60263 Cricetus
23	641	61.6	2331	6	AX365321 Sequence
24	641	61.6	2346	9	HSU8267 Homo sapien
25	634.5	61.0	2141	10	AF263239 Mus muscu
26	634.5	61.0	2198	10	BC013551 Mus muscu
27	633.5	60.9	720	6	AX365333 Sequence
28	633.5	60.9	720	9	AF176117 Homo sapi
29	632.5	60.8	597	6	AX365315 Sequence
30	632.5	60.8	597	10	AF237790 Mus muscu
31	628	60.3	2348	6	AX281651 Sequence
32	626.5	60.2	2212	6	AX365324 Sequence
33	626.5	60.2	2227	9	HSU85266 Homo sapien
34	626.5	60.2	2289	9	BC002864 Homo sapi
35	626.5	60.2	2407	9	AK092184 Homo sapi
36	615.5	59.1	621	10	AF282255 Mus muscu
37	615.5	59.1	2125	10	AF260717 Mus muscu
38	613.5	58.9	597	6	AX365312 Sequence
39	613.5	58.9	597	10	AF237789 Mus muscu
40	601	57.7	2173	6	AX410694 Sequence
41	601	57.7	2173	9	HSU28833 Homo sapien
42	600	57.6	2174	6	AR034241 Sequence
43	598	57.4	2284	9	HSU85265 Homo sapien
44	595	57.2	615	6	AX420432 Sequence
45	593	57.0	798	9	AF400429 Homo sapi

ALIGNMENTS

RESULT 1

LOCUS	AX365318	594 bp	DNA	linear	PAT 15-FEB-2002
DEFINITION	Sequence 8 from Patent WO0204491.				
ACCESSION	AX365318				
VERSION	AX365318.1	GI:18697047			
KEYWORDS					
SOURCE	house mouse.				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
AUTHORS	Williams, S.R. and Rochemme, B.				
TITLE	Methods and compositions relating to muscle selective calcineurin interacting protein (mcip)				
JOURNAL	Patent: WO 0204491-A 8 17-JAN-2002; Board of Regents, The University of Texas System (US) ; Williams, Sanders R. (US) ; Rochemme, Beverly (US)				
FEATURES	location/Qualifiers				
source	1..594				
	/organism="Mus musculus"				
	/db_xref="taxon:10090"				
CDS	1..594				
	/note="unnamed protein product"				
	/codon_start=1				
	/protein_id="CAD23810.1"				
	/db_xref="GI:18697048"				
	/translation="MPAPSMDCVSTLVACVVDVEVFTNOEYKEKFEGLFRTYDECVT FQLFKSFRVRVINFHPKPSAMARIL ELETORGRKKLKYFAQVOTPETDGKLIALA PQAPAFSLPPSSPVGWIKPILSDATPVNLDLYLAIVAKLIGGEKYEELHAGSTSPS VVHVDSMDSEEDPKTSKPKIOTRRRGLEPPSVSN"				
BASE COUNT	148 a 165 c 149 g 132 t				
ORIGIN					
Alignment Scores:					
Pred. No.:	2,49e-84	Length:	594		
Score:	1041.00	Matches:	197		
Percent Similarity:	100.00%	Conservative:	0		
Best Local Similarity:	100.00%	Mismatches:	0		
Query Match:	100.00%	Indels:	0		
DB:	6	Gaps:	0		
US-09-782-953-9 (1-197) x AX365318 (1-594)					
QY	1 MetProAlaProSerSerAspCyAspAlaValSerThrLeuValAlaCysValAlaAspVal	20			
DB	1 ATGCCAGCCCTGACATGACGTGATGATGTTCCACTCTGATCCGCTGTGTGGCGATGTC	60			
QY	21 GluValPheThrAngInGluValIyVgIuLysPheGluIyLeuPheArgThrTyAsp	40			
DB	61 GAGGCTTTTACCATGACGAGGAGGTTAAGAAATTCGAGGAGCATGTTCCGACCTATGAT	120			
QY	41 GluCysValThrPheGlnLeuPheLysSerPheArgValArgIleAspPheSerHis	60			
DB	121 GAATGCTGTGACGTTCCAGCGTGTAAAGAGTTTCGAGGGGTTGGAATTAATTCAGCCAT	180			
QY	61 ProLysSerAlaIaAlaArgAlaArgIleGlnLeuHisGlnThrGlnPheArgGlyIyLys	80			
DB	181 CCCAATCTCGACGCCCTGCGCCGATGACCTTCATAGACTCAGTTCACAGGAGGAAG	240			
QY	81 LeuLysLeuTyPheAlaGlnValGlnThrProGluThrAspGlyAspLysLeuHisLeu	100			
DB	241 CTAAACTCTACTTCGCCGCCAGGTCACAGCCCGACAGACAGATGAGACAAATGCAATTTG	300			
QY	101 AlaProProGlnProAlaLysGlnPheLeuIleSerProProSerSerProProValGly	120			
DB	301 GCACCTTCACAGCGCTGCGCAACAGTTCTCTATCTACACCCCTTCATATCTCTCTGGCC	360			
QY	121 TrpLysProIleSerAspAlaIaThrProValLeuAsnTyAspLeuLeuTyAlaValAla	140			
DB	361 TGGAGGCTTATCAGCGCATGGCACACGAGTCTCAACTATGACCTTTTATGCTGTGGC	420			
QY	141 LysLeuGlyProGlyGlyLysTyGluLeuHisAlaGlyThrGluSerThrProSerVal	160			

Db	421	AAACTAGGACAGGAGGAATATGAGCTGCACCGTCGGAACTGAGTCAACCGAGCGTC	480
Oy	161	VAlValHhSVAlCYsaPSeAspMeGlUGlUGlUglUuaPProlyThSerProlys	180
Db	481	GtGGcAcTgtGTGTgACAgCAGAcATGAGAGAGAGAGACCAGAACACTTCCCCNAAG	540
Oy	181	ProlYsIleIegInThArGArgProGIyleubPropoSeRalVseAsn	197
Db	541	CCMAAAATCATTCAGACCCGCCGTCGGGCTTGCCACCCTCCGTGTCCAAC	591
RESULT 2			
LOCUS	AF237791	594 bp	mRNA linear ROD 12-APR-2000
DEFINITION	Mus musculus myocyte-enriched calcineurin interactin protein 2		
ACCESSION	AF237791		
VERSION	AF237791.1	GI:7542530	
KEYWORDS			
SOURCE	Mus musculus.		
ORGANISM	Mus musculus		
REFERENCE	Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Eumalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 594) Rochemel,B.A., Vega,R.B., Yang,J., Wu,H., Baessel-Duby,R.S. and Williams,R.S. A Protein Encoded within the Down Syndrome Critical Region Is Enriched in Striated Muscles and Inhibits Calcineurin Signaling J. Biol. Chem. (2000) In press 2 (bases 1 to 594) Rochemel,B.A., Vega,R.B., Yang,J., Wu,H., Baessel-Duby,R.S. and Williams,R.S. Direct Submission Submitted (23-FEB-2000) Internal Medicine, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, TX 75390, USA		
TITLE	JOURNAL		
REFERENCE	AUTHORS		
CDS			
FEATURES	Bounce		
	1..594	/location=Qualifiers	
	/organism="Mus musculus"		
	/db_xref="taxon:10090"		
	1..594		
	/note="MCIP2; similar to Homo sapiens ZAK14; similar to EST sequence of GenBank Accession Number AU035927" /codon_start=1 /product="myocyte-enriched calcineurin interactin protein 2" /protein_id="AAF63487.1" /db_xref="gi:7542531" translation="MPAPSMDCDSTLVACYDVEVFPTNQEVKEFBLFRFYDECVL FQLSPRRIRINFSPKSLASRARLEHETOFRGKLCLYAQVOTPETDGKHLIALA PQPAPKPLISPPSSPVNGKSRIIDATPIELNLDLYAVAKLGEGRYELHAAGESTPPS VVHVCDSDMEEDBPDKTSFPKKITIGTRAPGLPRYSVN"		
BASE COUNT	148 a 165 c 149 g 132 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	2,496-84	Length:	594
Score:	1041.00	Matches:	197
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
Dbs:	10	Gaps:	0
US-09-782-953-9 (1-197) x AF237791 (1-594)			
Oy	1	MetProAlapRosemerAcAPCySaBPaVlaSerThrLeuValAlaCyVaValaaPval	20
Db	1	ATGCCAGCCCCTAGACATGAGACTGTGATTGTTCCACACTCTGTGCGCCGTGTGGTGATGTC	60
Oy	21	GUlValPherThraNgInglUvAllyVsGluLyepheGluGlyLeubPeaArgThrTyfASP	40
Db	61	GAGGCTCTTAACAACATCAGAGAGGTTTAGAAAAATTGAGGGAGCTGTTCCGGACCTATGAT	120
Oy	41	GUlCyValThrPheInleuPheLysserPheArgValArgIleAnPheSerHis	60

```

Db 121 GAATGTGTGACGCTTCCAGCTGTTAAGAGTTTCGACGGTTCGAATAAATTTTCAGCCAT 180
Qy 61 ProlysisSerAlaAaArgAlaArgileGluLeuHisGluThrGlnPheArgGlyLysLys 80
Db 181 CCCAAATCTGCAGCCCGTCCCGATAGAGCTTATGAGACTCAGTTCAGAGGGAAGAAG 240
Qy 81 LeuLysLeuTyrrPheAlaGlnValGlnThrProGluThrAspGlyAspLysLeuHisLeu 100
Db 241 CTAAAACTCTACTTTCGCCCGAGTCCAGACCCAGAGACAGATGAGACAAACTGCAATTTG 300
Qy 101 AlaProGlnProAlaLysGlnPheLeuLysSerProProSerSerProProValGly 120
Db 301 GCACCTCCAGAGCTTCCAAACAGTTCCTCATCTCACCCCTTCATCTCTCTCTCTCTCT 360
Qy 121 TrpLysProLysSerAspAlaThrProValLeuAsnTyrrAspLeuLeuTyrrAlaValAla 140
Db 361 TGAAGCCTATCAGGATGCCACACCATCTCACTATGACCTTCTTATGCTGTGGCC 420
Qy 141 LysLeuGlyProGlyGluLysTyrrGluLeuHisAlaGlyThrGluSerThrProSerVal 160
Db 421 AACTAGACACAGAGAGAAATATGAGCTGCACGCTGGAATCGAGTCTACACCGAGGTC 480
Qy 161 ValValHisValCysAspSerMetGluGluGluGluAspProLysThrSerProLys 180
Db 481 GTGGTGCATGTGTGACGCGCATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540
Qy 181 ProLysLeuLeuGlnThrArgArgProGlyLeuProProSerValSerAsn 197
Db 541 CCAAAAATCATTCAGACCCCGCGTCCGGCTTGCCACCTCCCTGTCACAC 591

RESULT 3
AF237887 659 bp mRNA linear ROD 17-NOV-2000
LOCUS Mus musculus Down syndrome candidate region 1-like protein 1
DEFINITION (Dscr11) mRNA, complete cds.
ACCESSION AF237887
VERSION AF237887.1 GI:7417330
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 659)
TITLES Strippoli, P., Petrini, M., Lenzi, L., Carinci, P. and Zannotti, M.
The murine DSCR1-like (Down syndrome candidate region 1) gene
family: conserved synteny with the human orthologous genes
JOURNAL Gene 257 (2), 223-232 (2000)
MEDLINE 20534792
PUBMED 11080588
REFERENCE 2 (bases 1 to 659)
AUTHORS Strippoli, P., Petrini, M., Lenzi, L., Carinci, P. and Zannotti, M.
TITLES Direct Submission
JOURNAL Submitted (22-FEB-2000) Istituto di Istologia ed Embriologia
Generale, Universita di Bologna, Via Belmeloro, 8, Bologna, BO
40126, Italy
FEATURES
Location/Qualifiers
1..659
/organism="Mus musculus"
/strain="BALB/c"
/db_xref="taxon:10090"
/chromosome="17"
/sex="male"
/tissue_type="brain"
1..659
/gene="Dscr11"
47..640
/gene="Dscr11"
/gene="Dscr11"
/note="similar to Homo sapiens Down syndrome candidate
region 1-like protein 1"
/codon_start=1
/product="Down syndrome candidate region 1-like protein 1"
/protein_id="AAF62538.1"

```

```

/db_xref="GI:7417331"
/translation="MPAPSMDCDVSTLVACVVDVEFTNQEVKEPEGLFRYDECVT
PQFMSFRVRINFSHPKSAARIEHETQFRGKGLYFAQVTPETDGGDLHAP
PQAKOFLISPPSPVGMKPIDATPVLYNDLLYAVAKLAGPGRKVELHAGTESTPSV
VHVDCSDMEEDPKTPKPKIIITRRPGLPPSVN"
BASE COUNT 167 a 183 c 163 g 146 t
ORIGIN
Alignment Scores: 2.79e-84 Length: 659
Pred. No.: 1041.00 Matches: 197
Score: 1041.00 Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 100.00% Gaps: 0
DB: 10
US-09-782-953-9 (1-197) x AF237887 (1-659)
Qy 1 MetProAlaProSerMetAspCysAspValSerThrLeuValAlaCysValValAspVal 20
Db 47 ATGCCAGCCCTAGCATGGACTGTGATGTTTCCACTCTGTCCTGTGTGTGGTGTGATGTG 106
Qy 21 GluValPheThrAsnGlnGluValLysGluLysPheGluGlyLeuPheArgThrTyrAsp 40
Db 107 GAGGTCTTTTACCAATCAGGAGGTAAAGGAAAAATTCAGGGACTGTTCGGGACCTATGAT 166
Qy 41 GluCysValThrPheGlnLeuPheLysSerPheArgValArgValArgIleAsnPheSerHis 60
Db 167 GAATGTGTGACGTTCCAGCTGTTTAAAGAGTTTCCACCGGTTTCGAATAAATTTACGCCAT 226
Qy 61 ProLysSerAlaAlaArgAlaArgileGluLeuHisGluThrGlnPheArgGlyLysLys 80
Db 227 CCCAAATCTGCAGCCCGTCCCGGATAGAGCTTCATGAGACTCAGTTCAGAGGGAAGAAG 286
Qy 81 LeuLysLeuTyrrPheAlaGlnValGlnThrProGluThrAspGlyAspLysLeuHisLeu 100
Db 287 CTAAAACTCTACTTCCCGCAGGTCCAGACCCAGAGACAGATGGAGACAAACTGCAATTTG 346
Qy 101 AlaProGlnProAlaLysGlnPheLeuLysSerProProSerSerProProValGly 120
Db 347 GCACCTCCAGACCTTCCAAACAGTTCCTCATCTCACCCCTTCATCTCTCTCTCTCTCT 406
Qy 121 TrpLysProLysSerAspAlaThrProValLeuAsnTyrrAspLeuLeuTyrrAlaValAla 140
Db 407 TGAAGCCTATCAGCGATGCCACACCATCTCTCACTATGACCTTCTTTATGCTGTGGCC 466
Qy 141 LysLeuGlyProGlyGluLysTyrrGluLeuHisAlaGlyThrGluSerThrProSerVal 160
Db 467 AAATAGGACACGAGGAGAAATATGAGCTGCACGCTGGAATCTTACACCGAGCGTC 526
Qy 161 ValValHisValCysAspSerMetGluGluGluGluAspProLysThrSerProLys 180
Db 527 GTGGTGCATGTGTGACAGCGACATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 586
Qy 181 ProLysLeuLeuGlnThrArgArgProGlyLeuProProSerValSerAsn 197
Db 587 CCAAAAATCATCCAGACCCCGCTCCCGGCTGCCACCTCCGCTGTCACAC 637

RESULT 4
AB061524 3227 bp mRNA linear ROD 13-NOV-2001
LOCUS Mus musculus mRNA for calcineurin inhibitory protein ZAKI-4,
DEFINITION complete cds.
ACCESSION AB061524
VERSION AB061524.1 GI:16904644
KEYWORDS
SOURCE Mus musculus cDNA to mRNA.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Kanou, Y., Miyazaki, T., Seo, H. and Murata, Y.
calcineurin inhibitory protein ZAKI-4

```

Journal Reference Authors Title	Journal	Features
Unpublished 2 (bases 1 to 3227)		
Kanou, Y., Miyazaki, T., Seo, H. and Murata, Y.		
Submitted (10-MAY-2001)		
Institute of Environmental Medicine, Furo-cho, Chikusa-ku, Aichi 464-8601, Japan (E-mail: kanou@iem.nagoya-u.ac.jp, Tel: 81-52-789-3888, Fax: 81-52-789-3876)		
Location/Qualifiers		
1..3227		
/organism="Mus musculus"		
/db_xref="taxon:10090"		
1..3227		
/gene="ZAKI-4"		
214..792		
/gene="ZAKI-4"		
/codon_start=1		
/product="calcineurin inhibitory protein ZAKI-4"		
/protein_id="BAB71955.1"		
/db_xref="gi:16904645"		
/translation="MDCVSTLVACVQVDFEVTNOVEKEEGFLRTYDECVTFLPSPRRIRINSPHKSAARARIELHETOPFGKKLYFAVOYFETGDKHLAPORAKFQSLPSRSPFHKRPKISDARITLKYDLILYAKLGRPEKVELHNGSTSTSVVNVQDSDMEEDDPKNSPKIITQTRRPLRPSVSN"		
BASE COUNT	905 a	715 c 742 g 865 t
ORIGIN		
Alignment Scores:		
Pred. No.:	2,82e-82	Length: 3227
Score:	1027.00	Matches: 194
Percent Similarity:	98.98%	Conservative: 1
Best Local Similarity:	98.48%	Mismatches: 2
Query Match:	98.66%	Indels: 0
DB:	10	Gaps: 0
US-09-782-953-9 (1-197) x AB061524 (1-3227)		
QY 1 MetProAlAProSerMetAaPcyAaPvalISerThrLeuValAlaCyValAlaPval	20	
DB 199 ATGCCAGCCCTGATGCATGCACTGGATGTTCCACTGTCGCGCTGTGTGGATGTG	258	
QY 21 GluValPheThrAngInGluValValGluValPheGluValPheArgThrTyAaP	40	
DB 259 GAGCTCTTACCAATCGAGGAGTTAGGAAATTCGAGGAGCTGTCGAGACTATGAT	318	
QY 41 GluCyValThrPheGluLeuPheLysSerPheArgValArgIleAsnPheSerHis	60	
DB 319 GAATGCTGACGTTCCACGCTGTTAGAGTTCCGAGGGTTGGAATATTCAGCAT	378	
QY 61 ProLysSerAlaAlaArgAlaArgIleGluLeuHisGluThrGlnPheArgGlyValVal	80	
DB 379 CCCAAATCTCAGCCGTCGCCGATGAGCTTCATGAGCTCAGTTCAAAGGAGAG	438	
QY 81 LeuLysLeuTyThrAlaGlnValGlnThrProGluThrArgGlyAaPValbLeuHisLeu	100	
DB 439 CTAAACTCTACTTGGCCAGGTCACAGCCCAAGACAGATGAGACAAATGCAATTG	498	
QY 101 AlaProGlnProAlaLysGlnPheLeuIleSerProProSerSerProProValGly	120	
DB 499 GCACTTCACAGCTGCGCAACAGATTCTCATCTCACCCCTTCATCTCCGCTGGG	558	
QY 121 TrpLysProIleSerAaPAlaThrProValLeuAsnTyAaPLeuLeuTyAlaValAla	140	
DB 559 TGGAGGCTTACAGCATGACACACAGCTCTCAATATGACCTCTTATAGCTGTG	618	
QY 141 LysLeuGlyProGlyGluValGlyGluLeuHisAlaGlyThrGlnSerThrProSerVal	160	
DB 619 AAATAGACACGAGAGAAATATGAGTGCACCTGAACTGAGTCTACACGAGGTC	678	
QY 161 ValValHisValCyAaPAspMetGluGluGluValPheProLysThrSerProLys	180	
DB 679 GTGGTGCATGCTGTGACAGCATGAGAGAGAGAGAGAGGACCCAAAGAAATTC	738	

QY	181	Prolysi1e1e1g1nThrArArProgl1yEuProPofSerValSerAsn	197
Db	739	CCNAAAATCATCCAGACCCCGCGTCCCGCCTGCACACCTCGTGTCCAC	789
RESULT 5			
LOCUS	AX329596	3184 bp	DNA
DEFINITION	Sequence 105 from Patent WO0194629.		linear
ACCESSION	AX329596		PAT 09-JAN-2002
VERSION	AX329596.1	GI:18102574	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,		
JOURNAL	Hortigan, S., Soppel, D.R. and Weaver, Z.		
FEATURES	Cancer gene determination and therapeutic screening using signature		
source	gene sets		
BASE COUNT	Patent: WO 0194629-A 105 13-DEC-2001;		
ORIGIN	Avalon Pharmaceuticals (US)		
Alignment Scores:	Location/Qualifiers		
Pred. No.:	1. 3184		
Score:	/Organism="Homo sapiens"		
Percent Similarity:	/db_xref="taxon:9606"		
Best Local Similarity:	98.98%		
Query Match:	96.45%		
	96.73%		
	6		
	Gaps:		
	0		
US-09-782-953-9 (1-197) x AX329596 (1-3184)			
QY	1	MetProAlaProSerMetAspCyAspValSerThrLeuValAlaCysValAlaAspVal	20
Db	190	ATCCAGCCCCCTGATGATGCTGTGATGTTTCCACCTCGTGTGCGTGTGATGTC	249
QY	21	GUUValPheThrAsnGlnGluValAspGluValPheGluGluLeuPheArgThrTyAsp	40
Db	250	GAGGCTCTTACCAATCAGGAGGTTAAGAAATTTGGGAGCTGTTTGGACTTTGAT	309
QY	41	GluCysValThrPheGlnPheLeuPheLeuSerPheArgValArgIleAsnPheSerHis	60
Db	310	GACTGTGTGACGTTCCAGCATTTTAAAGAGTTTCCAGCGTGTCCGTATTAACCTC	369
QY	61	ProlySerAlaAlaArgAlaArgIleGluLeuHisGluThrGlnPheArgGlyValys	80
Db	370	CCTAAATCTGCACCCGAGCTAGAAATAGAGCTTCATGAACCCAAATTCAGAGGAA	429
QY	81	LeuLysLeuThrPheAlaGlnValGlnThrProGluThrAspGlyAspLysLeuHisLeu	100
Db	430	TTAAAGCTTACTTTGCAAGCTTCAAGCTCCAGAGACAGATGGAGAAACAGCACTTG	489
QY	101	AlaProGlnProAlaLysGlnPheLeuLysSerProProSerSerProProValGly	120
Db	490	GCTCACCACCCAGCTGCGCAACAGTTTCTCATCTCCGCCCTTCTCCACCGCTTAC	549
QY	121	ThrLysProLysSerAspAlaThrProValLeuAsnTyAspLeuLeuTyAlaValAla	140
Db	550	TGGAGCCCATCAAGATGCGACGCCAGTCTCACTATGACCTTCTTATGCTGTGGCC	609
QY	141	LysLeuGlyProGlyGlyLysTyGluLeuHisAlaGlyThrGluSerThrProSerVal	160
Db	610	AAATTAAGACCGAGAGAGAAATATAGACTCAATGCAGGACTGAGTCCACCCCAAGTGC	669
QY	161	ValValHisValCysAspSerAspMetGluGluGluAspProLysThrSerProLys	180

Db 490 GCTCCACCCCGACCTGCGCAACAGTTTCTCATCTGCCCCCTTCTCCCGACCTGTAGC 549

Qy 121 TTPlySPProIleSerAspAlaThrProValLeuAanTYrAspLeuLeuTYrAlaValAla 140

 Db 550 TGGCAGCCCATCAAGATGCCAGCGACCTCTCACTATGACTCTCTATCTGTGGCC 609

Qy 141 LysLeuGIyProGIyGIuLySTyrgIuLeuHIsAlaGIyThrGIuSerThrProSerVal 160

 Db 610 AAACTAGGACCAAGAGAGATATAGCTCCATGCGAGGAGCTGAGTCCACCCCAAGTGC 669

Qy 161 ValValHisValAlCyAspSerAspMetGIuGIuGIuGIuAspProLyThrSerProLyS 180

 Db 670 GTCTGACACGTGTGGACAGTACATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 729

Qy 181 ProLySIIleIleGIInThrArgArgProGIyLeuProProSerValSerAsn 197

 Db 730 CCAAAATATATCAAACTGGCGTCTGAGCTGCCACCTCCGTGTCCAAC 780

RESULT 8
 AX420436 3184 bp DNA linear PAT 18-JUN-2002
 LOCUS AX420436
 DEFINITION Sequence 12 from Parent W00204513.
 ACCESSION AX420436
 VERSION AX420436.1 GI:21524589
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS 1
 TITLE Loring,J.F., Tingley,D.W., Edwards,C.M. and Streeter,D.G.
 JOURNALS Down syndrome critical region 1-like 1 proteins
 Patent: WO 0204513-A 12 17-JAN-2002;
 Incyte Genomics, Inc. (US)
 FEATURES
 source 1..3184
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /note="Incyte ID No: 91435039"
 BASE COUNT 921 a 681 c 657 g 925 t
 ORIGIN

Alignment Scores:
 Pred. No.: 1..78-80 Length: 3184
 Score: 1007.00 Matches: 190
 Percent Similarity: 98.98% Conservative: 5
 Best Local Similarity: 96.45% Mismatches: 2
 Query Match: 96.73% Indels: 0
 DB: Gaps: 0

US-09-782-953-9 (1-197) x AX420436 (1-3184)

Qy 1 MetProAlaProSerMetAspCyAspValSerThrLeuValAlaCyValValAlaVal 20

 Db 190 ATGCCAGCCCTAGCAGTACGTGTATGTTCCACCTCGTTCCTGTGGGTGATGTC 249

Qy 21 GIuValPheThrAsnGIInGIuValLySGIuLyPheGIuGIyLeuPheArgThrTYrAsp 40

 Db 250 GAGGTCTTACCAATAGAGAGGTTAGAGAAATTTGGGGAGCTGTTCGACTTATGAT 309

Qy 41 GIuCyValAlaThrPheGIInLeuPheLySerPheArgArgValArgIleAsnPheSerHis 60

 Db 310 GACTGTGTACAGTTCACGATTTTAAGAGTTTCAGACGTGTCTGTAAACCTCAGCAAT 369

Qy 61 ProLySerAlaAlaArgAlaArgIleGIuLeuHIsGIInThrGIInPheArgGIyLyLyS 80

 Db 370 CCTAAATCTGACGCCGAGCTAGATAGAGCTTCATGAAACCAATTCAGAGGAGAAAAA 429

Qy 81 LeuLyLeuTYrPheAlaGIInValGIInThrProGIuThrAspGIyAspLySLeuHIsLeu 100

 Db 430 TTAAAGCTCTACTTTCACAGGTTCCAGACTCCAGAGACAGATGAGACAAACTGCACTTG 489

Qy 101 AlaProGIInProAlaLyGIInPheLeuIleSerProProSerSerProProValGIy 120

 Db 490 GCTCCACCCCGACCTGCGCAACAGTTTCTCATCTGCCCCCTTCTCCCGACCTGTAGC 549

Qy 121 TTPlySPProIleSerAspAlaThrProValLeuAanTYrAspLeuLeuTYrAlaValAla 140

 Db 550 TGGCAGCCCATCAAGATGCCAGCGACCTCTCACTATGACTCTCTATCTGTGGCC 609

Qy 141 LysLeuGIyProGIyGIuLySTyrgIuLeuHIsAlaGIyThrGIuSerThrProSerVal 160

 Db 610 AAACTAGGACCAAGAGAGATATAGCTCCATGCGAGGAGCTGAGTCCACCCCAAGTGC 669

Qy 161 ValValHisValAlCyAspSerAspMetGIuGIuGIuGIuAspProLyThrSerProLyS 180

 Db 670 GTCTGACACGTGTGGACAGTACATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 729

Qy 181 ProLySIIleIleGIInThrArgArgProGIyLeuProProSerValSerAsn 197

 Db 730 CCAAAATATATCAAACTGGCGTCTGAGCTGCCACCTCCGTGTCCAAC 780

RESULT 9
 HUMZAKI4 3184 bp mRNA linear PRI 06-FEB-1999
 LOCUS HUMZAKI4
 DEFINITION ZAKI-4 mRNA in human skin fibroblast, complete cds.
 ACCESSION D83407.1
 VERSION D83407.1 GI:1435039
 KEYWORDS ZAKI-4; thyroid hormone responsive.
 SOURCE Homo sapiens skin fibroblast cDNA to mRNA.
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS 1
 TITLE Miyazaki,T., Kanou,Y., Murata,Y., Ohmori,S., Niwa,T., Maeda,K.,
 Yamamura,H. and Seo,H.
 JOURNALS Molecular cloning of a novel thyroid hormone responsive gene,
 ZAKI-4, in human skin fibroblasts
 J. Biol. Chem. 271 (24), 14567-14571 (1996)
 MEDLINE 96278928
 REFERENCE 2 (bases 1 to 3184)
 AUTHORS Miyazaki,T.
 JOURNALS Unpublished
 REFERENCE 3 (bases 1 to 3184)
 AUTHORS Miyazaki,T.
 JOURNALS Direct Submission
 Submitted (06-FEB-1996) Takashi Miyazaki, Research Institute of
 Environmental Medicine, Nagoya Univ., Department of Endocrinology
 and Metabolism; Furo-chou, Chikusa-ku, Nagoya, Aichi 464-01, Japan
 (E-mail:tamawar@em.nagoya-u.ac.jp, Tel:052-789-3867,
 Fax:052-789-3887)

FEATURES
 source 1..3184
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /cell_type="skin fibroblast"

gene 1..3184
 /gene="ZAKI-4"
 205..783
 /note="A thyroid hormone responsive gene in human skin
 fibroblasts"

CDS
 1..3184
 /protein_id="BA011911.1"
 /codon_start=1
 /translation="MDCDVSTLVACVVDVEFTVQEVKEKFGGLFRTYDDCVTFQLFK
 SFRVRINFSNPKSARARIELHETGRGKLLKYPVQVOTPETDGLHLAPQPAK
 OFLSPSPSPSPVSNQPIINDATPVNTDILAVVALGGEKELEHAGRESTSVVYHVC
 DSDLEEDDEPKTSFKPIIDTRRGLPSPVSN"
 1023..1028
 /gene="ZAKI-4"

polyA_signal
 1044
 /gene="ZAKI-4"
 3168..3173
 polyA_signal
 /gene="ZAKI-4"

BASE COUNT	921 a	681 c	657 g	925 t
ORIGIN				

Alignment Scores:

Pred. No.:	1.7e-80	Length:	3184
Score:	1007.00	Matches:	190
Percent Similarity:	98.98%	Conservative:	5
Best Local Similarity:	96.45%	Mismatches:	2
Query Match:	96.73%	Indels:	0
DB:	9	Gaps:	0

US-09-782-953-9 (1-197) x HUMZAKI4 (1-3184)

QY	1	Met	Pro	Ala	Pro	Ser	Met	Asp	Cys	Asp	Val	Ser	Thr	Leu	Val	Ala	Cys	Val	Val	Asp	Val	20		
DB	190	ATGC	CAG	CCCT	TAG	CA	TGG	ACT	GTG	TAT	TTT	TCA	CT	TGG	CT	TGG	CT	TGG	TGG	TAT	GC	249		
QY	21	Glu	Val	Phe	Thr	Asn	Gln	Glu	Val	Leu	Glu	Leu	Phe	Glu	Gly	Leu	Phe	Asn	Thr	Tyr	Asp	40		
DB	250	GAGT	CTTT	ACCA	TAT	CAG	GAG	TTT	AAG	AAA	AAA	TTT	TGG	GGA	CT	TGT	TTC	GGA	CTT	TAT	GAT	309		
QY	41	Glu	Cys	Val	Thr	Phe	Gln	Leu	Phe	Leu	Ser	Phe	Arg	Arg	Val	Arg	Gln	Leu	Asn	Phe	Ser	His	60	
DB	310	GACT	TGT	GAG	CTT	TCC	AGC	TAT	TTA	AGAG	TTT	CAG	AG	CTT	CAG	AG	CTT	CGG	TAT	AAAT	CTT	CAG	CAAT	369
QY	61	Pro	Lys	Ser	Ala	Ala	Arg	Ala	Arg	Ile	Glu	Leu	His	Glu	Thr	Gln	Phe	Arg	Gly	Lys	Lys	80		
DB	370	CCT	AAAT	CTC	GAG	CCC	GAG	CT	TAG	GAT	AG	AG	CTT	CAT	GAA	ACC	CCAA	TT	CAG	AGG	GAA	AAAA	429	
QY	81	Leu	Lys	Leu	Tyr	Phe	Ala	Gln	Val	Gln	Thr	Pro	Glu	Thr	Asp	Gly	Asp	Lys	Leu	His	Leu	100		
DB	430	TTA	AGT	CT	TAC	TTT	G	CAC	AGG	TT	CAC	ACT	T	CAG	AC	T	T	CAG	AC	T	T	489		
QY	101	Ala	Pro	Pro	Gln	Pro	Ala	Lys	Gln	Phe	Leu	Ile	Ser	Pro	Pro	Ser	Ser	Pro	Pro	Val	Gly	120		
DB	490	GCT	CC	ACCC	CAG	CTG	CC	AA	CAG	TTT	CT	CAT	CT	CG	CC	CC	CT	TCT	CT	CC	CA	CT	549	
QY	121	Trp	Lys	Pro	Ile	Ser	Asp	Ala	Thr	Pro	Val	Leu	Asn	Tyr	Asp	Leu	Leu	Tyr	Ala	Val	Ala	140		
DB	550	TGG	CAG	CCCAT	CA	CA	AGT	ATG	CA	GC	CA	GC	CA	AGT	CT	CT	CA	ACT	TAT	GAC	CT	CT	609	
QY	141	Lys	Leu	Gly	Pro	Gly	Glu	Lys	Tyr	Glu	Leu	His	Ala	Gly	Thr	Glu	Ser	Thr	Pro	Ser	Val	160		
DB	610	AA	CT	TAG	CA	GAG	AGA	AGT	AT	GAG	CT	CA	TG	CA	GGA	GG	ACT	GAG	TCC	AC	CC	CA	669	
QY	161	Val	Val	His	Val	Cys	Asp	Ser	Asp	Met	Glu	Glu	Glu	Asp	Pro	Lys	Thr	Ser	Pro	Lys	180			
DB	670	GTC	TG	CA	CGT	GTG	CG	CA	CT	GAC	AT	GAC	AT	GAC	CT	CA	CT	CA	CT	CA	CT	729		
QY	181	Pro	Lys	Ile	Ile	Gln	Thr	Arg	Arg	Pro	Gly	Leu	Pro	Ser	Val	Ser	Asn	197						
DB	730	CC	AAAA	TAT	CA	CA	ACT	CG	CGT	CT	CG	CGT	CT	CG	CGT	CT	CG	CGT	CT	CG	CGT	780		

RESULT 10					
AB061525	AB061525	3295 bp	muscle mRNA	linear	ROD 13-NOV-2001
LOCUS	Mus musculus mRNA for calcineurin inhibitory protein ZAKI-4 beta,				
DEFINITION	complete cds.				
ACCESSION	AB061525				
VERSION	AB061525.1	GI:16904646			

1 mammalian; eutheria; rodentia; sciurognathini

REFERENCE
2 (bases 1 to 3295)
AUTHORS Kanou, Y., Miyazaki, T., Seo, H. and Murata, Y.
TITLE Direct Submission

Institute of Environmental Medicine; Furo-cho, Chikusa-ku, Nagoya,
Aichi 464-8601, Japan (E-mail: kanou@iem.nagoya-u.ac.jp,
Tel: 81-52-789-3888, Fax: 81-52-789-3876)

FEATURES
source

```

gene
  /organism="Mus musculus"
  /db_xref="taxon:10090"
  1..3295
  /gene="ZAKI-4 beta"
CDS
  116..847
  /gene="ZAKI-4 beta"
  /codon_start=1

```

BASE COUNT	939 a	714 c	759 g	879 t	4 others
ORIGIN					

Alignment Scores:

Pred. No.:	3.02e-73	Length:	3295
Score:	926.00	Matches:	176
Percent Similarity:	95.72%	Conservative:	3
Best Local Similarity:	94.12%	Mismatches:	8
Query Match:	88.95%	Indels:	0
DB:	10	Gaps:	0

US-09-782-953-9 (1-197) X AB061525 (1-3295)

Qy	11	SerThrLeuValAlaCysValValAspValGluValPheThrAsnGlnGluValLysGlu	30
Db	284	AACTCATATGTTTCCCTGCATGTTCACCACTCTGTGTTTGAAGAAGAGGAGACAAGAA	343
Qy	31	LysPheGluGlyLeuPheArgThrTyrAspGluCysValThrPheGlnLeuPheLysSer	50
Db	344	AAATTTGAGGGACTGTTCCGACCTATGATGAATGTGTGACGTTCAGCTGTTTAAAGT	403
Qy	51	PheArgArgValArgIleAsnPheSerHisProLysSerAlaAlaArgAlaArgIleGlu	70
Db	404	TTCCGACGGGTTCCGATTAATTTTCAGCCATCCCAATCTGCAGCCGTGCCCGGATAGAG	463
Qy	71	LeuHisGluThrGlnPheArgGlyLysLysLeuLysLeuTyrPheAlaGlnValGlnThr	90
Db	464	CTTTCATGAGACTCAGTTCAGAGGAAAGAGCTAAAACTCTACTTCGCCCAGGTCCAGACC	523
Qy	91	ProGluThrAspGlyAspLysLeuHisLeuAlaProProGlnProAlaLysGlnPheLeu	110
Db	524	CCAGACAGATGGAGACAACTGCATTTTGGACCTTCACAGCTCCCAACAGTTCCCTC	583
Qy	111	IleSerProProSerSerProProValGlyTyrLysProIleSerAspAlaThrProVal	130
Db	584	ATCTCACCCCTTCATCTCTCTCTGTGGCTGGAGCCCTATCAGCGATGCCACACCAGTC	643
Qy	131	LeuAsnTyrAspLeuLeuTyrAlaValAlaLysLeuGlyProGlyGlnLysTyrGluLeu	150
Db	644	CTCAACTATGACCTCTCTTATGCTGTGGCCAAACTAGGACCAGGAGAGAAATATGAGCTG	703
Qy	151	HisAlaGlyThrGluSerThrProSerValValHisValCysAspSerAspMetGlu	170
Db	704	CACGCTGGAACTGAGTCTACCGAGCGTGTGGTCATGTGTGACAGCGACATGGAG	763
Qy	171	GluGluGluAspProLysThrSerProLysProLysIleIleGlnThrArgArgProGly	190
Db	764	GAGGAGGAGGACCCAAAGACTTCCCCAAAGCCAAATCATCCAGACCCGCGCTCCCGGC	823
Qy	191	LeuProProSerValSerAsn	197
Db	824	CTGCCACCCCTCGTGTCCAAC	844

```

RESULT 11
AX420427/c 1021 bp DNA linear PAT 18-JUN-2002
LOCUS
DEFINITION Sequence 3 from Patent WO0204513.
ACCESSION AX420427
VERSION AX420427.1 GI:21524582
KEYWORDS
SOURCE
ORGANISM
Norway rat.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1
AUTHORS Loring,J.F., Tingley,D.W., Edwards,C.M. and Screeter,D.G.
TITLE Down syndrome critical region 1-like 1 protein
JOURNML Patent: WO 0204513-A 3 17-JAN-2002;
Incyte Genomics, Inc. (US)
FEATURES
Location/Qualifiers
source
1..1021
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/notice="Incyte ID No: 219435_Rn.1"
BASE COUNT 237 a 241 c 286 g 257 t
ORIGIN

Alignment Scores:
Pred. No.: 2,266-72 Length: 1021
Score: 910.00 Matches: 192
Percent Similarity: 94.63% Conservative: 2
Best Local Similarity: 93.66% Mismatches: 3
Query Match: 87.42% Indels: 8
DB: Gaps: 0

US-09-782-953-9 (1-197) x AX420427 (1-1021)
QY 1 MetProAlaPheSerMetLeuPheValSerThrLeuValAlaCysValValAspVal 20
Db 855 ATGCCAGCCCTAGCATGATGATGATGATGATGATGATGATGATGATGATGATG 796
QY 21 GluValPheThrAsnGlnGluValLysGluLysPheGluGlyLeuPheArgThrTyra 40
Db 795 GAGGCTTTACCAATCAGAGGTTTAAAGAAATTTGAGGAGCTGTTCCGGAACCTTAC 736
QY 41 GluCysValThrPheGlnLeuPheLysSerPheArgArgValArg-IleAsnPheSer 60
Db 735 GATGTGTGACGTTTACGTTTAAAGATTTCCGAGGTTTGGAAATTAATTTACGCA 676
QY 60 sProLysSerAlaAlaArg-AlaArgIle-GluLeuHisGlu-ThrGlnPheArg-Gly 79
Db 675 CCCCAAGCTGACGAGCCGCTGCCGATGAGAGCTTCATGTGAGACCACTTATGAGGA 616
QY 79 ValValLeuLysLeuTyrPheAlaGlnValGlnThrProGlnThrAspGly-AspLysLeu 98
Db 615 ACAAAGCTGAACCTTACTTTGACACAGTCCAGACCCCAAGACAGATGTGATGACAAACTG 556
QY 99 His-LeuAlaProProGlnProAlaLysGlnPheLeuLieserProPheSerSerPro 118
Db 555 CAGTTTGGACACCCCAACACCTGCAACAGTTCTCATCTCACCCCTTCACTCTCTCC 496
QY 118 oValGlyTyrLysProLieserAspAlaThrProValLeuAsn-TyrAspLeuLeuTyr 138
Db 495 CCTTGCTGGAAGCCTATCAGCATGCCACACAGTCTCAACGTAAGACCTCTTATG 436
QY 138 lAlaValAlaLysLeuGlyProGlyGlyLysTyrGlnLeuHisAlaGlyThrGlnSerThr 158
Db 435 CCCTGGCCCAACTAGACAGACAGAGAAATATGAGCTGATGCGGGAAGTGAATCAGC 376
QY 158 rSerSerValValValHisValCysAspSerAspMetGluGluGluGluAspProLysTh 178
Db 375 CAGAGCTTTCTGTCACGCTGTGTGACAGCAGCTTGGAGAGAGAGATCCAAACACTT 316
QY 178 eProLysProLysIleIleGlnThrArgArgProGlyLeuProProSerValSerAsn 197

```

```

Db 315 CCCCAAGCCAAAATATATCCAGACCCGCGCTCTGCGCTCCACCTCCGTTCAC 257
RESULT 12
AX074351 934 bp DNA linear PAT 06-FEB-2001
LOCUS
DEFINITION Sequence 3 from Patent WO0104305.
ACCESSION AX074351
VERSION AX074351.1 GI:12710514
KEYWORDS
SOURCE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1
AUTHORS Tang,Y.T. and Yue,H.
TITLE Human proteins involved in detoxification
JOURNML Patent: WO 0104305-A 3 18-JAN-2001;
Incyte Genomics, Inc. (US)
FEATURES
Location/Qualifiers
source
1..934
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 247 a 244 c 206 g 237 t
ORIGIN

Alignment Scores:
Pred. No.: 5,732-72 Length: 934
Score: 905.00 Matches: 171
Percent Similarity: 95.72% Conservative: 8
Best Local Similarity: 91.44% Mismatches: 8
Query Match: 86.94% Indels: 0
DB: Gaps: 0

US-09-782-953-9 (1-197) x AX074351 (1-934)
QY 11 SerThrLeuValAlaCysValValAspValGluValPheThrAsnGlnGluValLysGlu 30
Db 161 AACTCGTGTGTGCTGCGTGAATGTCACAGTCAGTGTGGAAGAGAGAGCAAGCAAA 220
QY 31 LysPheGlnGlyLeuPheArgThrTyraAspGluCysValThrPheGlnLeuPheLysSer 50
Db 221 AATTTGAGGAGCTGTTCCGACTTATGATGATGATGATGATGATGATGATGATGAT 280
QY 51 PheArgArgValArgIleAsnPheSerHisProLysSerAlaAlaArgAlaArgIleGlu 70
Db 281 TTCACAGCTGCCGATTAACCTTCAGCAATCTTAATCTGCAGCCGAGTAGATAGAG 340
QY 71 LeuHisGlnThrGlnPheArgGlyLysLeuLysLeuLysLeuTyrPheAlaGlnValGlnThr 90
Db 341 CTTCAATGAAACCAATTCAGAGGGAATAAATTAAGCTCTACTTGGACAGGTTCCAGACT 400
QY 91 ProGlnThrArgGlyAspLysLeuHisLeuAlaProProGlnProAlaLysGlnPheLeu 110
Db 401 CCAGAGACAGATGAGAGCAAACTGCACTTGAGCTCCACCCAGCTGCCAAACAGTTTCTC 460
QY 111 LieserProProSerSerProProValGlyTyrLysProLieserAspAlaThrProVal 130
Db 461 ATCTGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 520
QY 131 LeuAsnTyrAspLeuLeuTyrAlaValAlaLysLeuGlyProGlyGlyLysTyrGlnLeu 150
Db 521 CTCACACTATGACCTCTCTCTATGCTGTGCGCAAACTAGAGACAGAGAGAGATATAGCTC 580
QY 151 HisAlaGlyThrGlnSerThrProSerValValValHisValCysAspSerAspMetGlu 170
Db 581 CATGAGGAGACTGATCCACCCCAAGTGTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 640
QY 171 GluGlnGluAspProLysThrSerProLysProLysIleIleGlnThrArgArgProGly 190
Db 641 GAAAGAGAGAGACCAAAAGACTTCCCAAGCCAAAATCATCTCAAACTCGGAGTCTCTGC 700
QY 191 LeuProProSerValSerAsn 197

```



```

Db 701 *CTGCCACCTCCGTGTCCAAC 721
RESULT 13
AX420425
LOCUS
DEFINITION Sequence 1 from Patent WO204513.
ACCESSION AX420425
VERSION AX420425.1 GI:21524581
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Loring, J.F., Tingley, D.W., Edwards, C.M. and Streeter, D.G.
TITLE Down syndrome critical region 1-like 1 proteins
JOURNAL Patent: WO 0204513-A 1 17-JAN-2002;
Incyte Genomics, Inc. (US)
FEATURES
source
Location/Qualifiers
1..3159
/organism="Homo sapiens"
/db_xref="taxon:9606"
/notes="Incyte ID No: 247500.5"
BASE COUNT 926 a 657 c 657 g 915 t 4 others
ORIGIN

Alignment Scores:
Pred. No.: 2,17e-71 Length: 3159
Score: 905.00 Matches: 171
Percent Similarity: 95.72% Conservative: 8
Best Local Similarity: 91.44% Mismatches: 8
Query Match: 86.94% Indels: 0
DB: 6 Gaps: 0

US-09-782-953-9 (1-197) x AX420425 (1-3159)

QY 11 SerThrLeuValAlaCysValValAspValGluValPheThrAsnGlnGluValLysGlu 30
Db 389 AACTCGTGTTCGGTGCATGTTCCAGCTGATGATGCTGTGTGAGCTTCAGCTATTTAAGAGT 508
QY 31 LysPheGluGlyLeuPheArgThrTyrAspGluCysValThrPheGlnLeuPheLysSer 50
Db 449 AAATTTGAGGAGCTGTTTCGGACTTATGATGACTGTGTGAGCTTCAGCTATTTAAGAGT 508
QY 51 PheArgArgValArgLeuAsnPheSerHisProLysSerAlaAlaArgAlaArgIleGlu 70
Db 509 TTCAGAGCTGTCGGTATATAACTTTCAGCAATCTTAATCTGCAGCCCGAGCTAGGATAGAG 568
QY 71 LeuHisGluThrGlnPheArgGlyLysLeuLysLeuLysLeuLysLeuLysLeuLysLeu 90
Db 569 CTTCAATGAAACCCCAATTCAGAGGAAAAAATTAAAGCTCTACTTTGCACAGGTTTCAGACT 628
QY 91 ProGluThrAspGlyAspLysLeuHisLeuAlaProGlnProAlaLysGlnPheLeu 110
Db 629 CCAGAGACAGATGGACAACTGCATTCGCTCCACCCAGAGCTGCCAAACAGTTTCTC 688
QY 111 IleSerProProSerSerProProValGlyTyrPheProLysSerAspAlaThrProVal 130
Db 689 ATCTGCGCCCTCTCTCTCCACCTGTTGGCTGGCAGCCCATCAACGATGCCAGCCAGTC 748
QY 131 LeuAsnTyrAspLeuLeuTyrAlaValAlaLysLeuGlyProGlyGluLysTyrGluLeu 150
Db 749 CTCACTATGACCTCTCTATGCTGTGGCCAACTAGGACGAGAGAGATATAGACTC 808
QY 151 HisAlaGlyThrGluSerThrProSerValValHisValCysAspSerAspMetGlu 170
Db 809 CATGCAGGAGCTGAGTCCACCCCAAGTGTGCTGCACGCTGCGACAGTGCATAGAG 868
QY 171 GlyGluAspProLysThrSerProLysProLysIleIleGlnThrArgArgProGly 190
Db 869 GAAGAAAGAGACCCAAAGACTTCCCAAGCCAAAAATCATCAAACTCGGCGCTCTGGC 928
QY 191 LeuProProSerValSerAsn 197

```

```

|||||
Db 929 CTGCCACCTCCGTGTCCAAC 949
RESULT 14
AK090990
LOCUS
DEFINITION Homo sapiens cDNA FLJ33671 fis, clone BRAWH2001459, highly similar
to ZAKI-4 PROTEIN.
ACCESSION AK090990
VERSION AK090990.1 GI:21749260
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens brain cDNA to mRNA, clone_lib:BRAWH2
clone:BRAWH2001459.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Nishi, T., Ota, T., Nakagawa, S., Senoh, A., Mizuguchi, H., Inagaki, H.,
Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S.,
Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T.,
Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M.,
Kikuchi, H., Kanda, K., Wagatsuma, M., Murakawa, K., Kanehori, K.,
Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y.,
Sugano, S., Nagahari, K., Masuho, Y., Nagai, K. and Isogai, T.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3261)
AUTHORS Isogai, T. and Yamamoto, J.
TITLE Direct Submission
JOURNAL Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing;
Research Association for Biotechnology (RAB); CDNA library
construction; Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing; RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing; HRI and
RAB; annotation: HRI and RAB.
FEATURES
source
Location/Qualifiers
1..3261
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="BRAWH2001459"
/tissue_type="brain"
/clone_lib="BRAWH2"
/notes="cloning vector: pME18SFL3"
BASE COUNT 943 a 710 c 679 g 929 t
ORIGIN

Alignment Scores:
Pred. No.: 2,24e-71 Length: 3261
Score: 905.00 Matches: 171
Percent Similarity: 95.72% Conservative: 8
Best Local Similarity: 91.44% Mismatches: 8
Query Match: 86.94% Indels: 0
DB: 9 Gaps: 0

US-09-782-953-9 (1-197) x AK090990 (1-3261)

QY 11 SerThrLeuValAlaCysValValAspValGluValPheThrAsnGlnGluValLysGlu 30
Db 293 AACTCGTGTTCGGTGCATGTTCCAGCTGATGATGCTGTGTGAGCTTCAGCTATTTAAGAGT 352
QY 31 LysPheGluGlyLeuPheArgThrTyrAspGluCysValThrPheGlnLeuPheLysSer 50
Db 353 AAAATTTGAGGAGCTGTTTCGGACTTATGATGATGCTGTGACGTTCCAGCTATTTAAGAGT 412
QY 51 PheArgArgValArgIleAsnPheSerHisProLysSerAlaAlaArgAlaArgIleGlu 70
Db 413 TTCAGACGTCGTCGTATTAACCTTCAGCAATCTCAATCTGCAGCCCGAGCTAGGATAGAG 472

```

polyA_signal	1079..1084	DB3407"
/note="alternative"		
polyA_signal	3224..3229	
BASE COUNT	947 a 688 c 677 g 928 t	
ORIGIN		

Alignment Scores:		
Pred. No.:	3.23e-70	Length: 3240
Score:	892.00	Matches: 169
Percent Similarity:	94.65%	Conservative: 8
Best Local Similarity:	90.37%	Mismatches: 10
Query Match:	85.69%	Indels: 0
DB:	9	Gaps: 0

US-09-782-953-9 (1-197) x AY034085 (1-3240)		
Qy	11 SerThrLeuValAlaCysValAlaAspValGluValPheThrAsnGluValLysGlu	30
Db	276 AACGCGTTGTTGGCTGCATNTTCCACCACTCACTGTTTAAAGAGAGAAAGACCAAGAA	335
Qy	31 LysPheGluGlyLeuPheAsnGlyThrTyraSPGluCysValThrPheGluLeuPheLysSer	50
Db	336 AAATTTGGGGGACGTTTCCGACCTTAAGAAGACTGTGTGACGTTCCAGCACTTAATTAAGAGT	395
Qy	51 PheArgArgValArgGluLeuAsnPheSerHisProLysSerAlaAlaArgAlaArgGlu	70
Db	396 TTCAAGACGTCGCTATTAATCTTACGAAATCTTAATCTCAAGCCGAGCTTAGATTAAG	455
Qy	71 LeuHisGluThrGlnPheAsnGlyLysLysLeuLysLeuLysLeuTyraPheAlaGluValGlnThr	90
Db	456 CTTCAATGAACCAATTCAGAGGGGAAAAATTAAAGCTCACTTGTGCACAGTTCAAGACT	515
Qy	91 ProGluThrAspGlyAspLysLeuHisLeuAlaProProGluInProAlaLysGlnPheLeu	110
Db	516 CCAGGACAGATGAGAGCAAACTGCACCTGGCTCCACCCAGCCTGCCAAACAGTTTCTC	575
Qy	111 IleSerProProSerSerProProValGlyTyrLysProIleSerAspAlaThrProVal	130
Db	576 ATCTGCCCCCTTCTCCCTCCCACTGTTAGCTGGAGGCCCATCAACGATGCGACGCCAGTC	635
Qy	131 LeuAsnTyraSPLeuLeuTyraValAlaAlaLysLeuGlyProGlyGluLysTyraGluLeu	150
Db	636 CTCCAACTATGACCTCTCTATAGCTGTGGCCAAACTAGACACAGAGAAAGTATGACCTC	695
Qy	151 HisAlaGlyThrGluSerThrProSerValValAlaHisValCysAspSerAspMetGlu	170
Db	696 CATGAGGAGACTGAGTCCACCCCAAGTGTGTCGTGCACGTTGCGACAGTGCATATGAG	755
Qy	171 GluGluGluAspProLysThrSerProLysProLysAlaIleGlnThrArgArgProGly	190
Db	756 GAAGAAAGGAGCCCAAAAGACTTCCCAAAAGCCAAAAATCATCCAACTCGGCGTCTGGC	815
Qy	191 LeuProProSerValSerAsn	197
Db	816 CTGCCACCTCCGTTGCCAAC	836

Search completed: December 14, 2002, 21:32:11
Job time : 1654.42 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 14, 2002, 19:56:01 ; Search time 1112.9 Seconds
(without alignments)
2866.843 Million cell updates/sec

Title: US-09-782-953-9

Perfect score: 1041

Sequence: 1 MFAPMDCVSTLVACVDV.....SPKPKIQRRLPGLPPSVSN 197

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-Q/cgn2_1/USPTO spool/US09782953/runat_11122002_114429_17353/app query.fasta_1.1173
-DB=EST -OFMT=fasta -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09782953 @CNC 1 1 1583 @runat_11122002_114429_17353 -ICPU=3
-NO_XLPXY -NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1020	98.0	925	14	BQ720798
2	1007	96.7	1055	14	BM925558
3	994	95.5	903	9	AL532776
4	941.5	90.4	847	13	BI545609
5	854	82.0	753	10	AW957141
6	849	81.6	758	12	BG709000
7	805	77.3	495	12	BF443461
8	803	77.1	742	12	BG721994
9	788	75.7	546	14	BQ260356
10	786.5	75.6	617	13	BI476888
11	774	74.4	640	10	BB650331
12	758.5	72.9	1007	12	BF534085
13	748	71.9	557	9	AI796134
14	728	69.9	626	12	BG701025
15	726	69.7	747	9	AI69846
16	715	68.7	516	9	AL693037
17	674.5	64.8	1103	11	AK019377
18	669	64.3	530	9	AA814235
19	658.5	63.3	863	14	BQ438012
20	655.5	62.8	829	9	AU035927
21	646	62.1	550	12	BG80881
22	641	61.6	837	9	AU124628
23	641	61.6	931	9	AL543576
24	641	61.6	939	9	AL546617
25	641	61.6	1041	13	BM450020
26	636	61.1	946	14	BQ278576
27	635	61.0	718	9	AU131040
28	634	60.9	662	13	BI464521
29	630.5	60.6	1158	14	BQ895506
30	630	60.5	642	14	BQ520533
31	630	60.5	689	9	AL652662
32	630	60.5	748	14	BM964376
33	628	60.3	433	9	AA248590
34	628	60.3	885	9	AL544755
35	626.5	60.2	828	9	AL536447
36	626.5	60.2	867	14	BQ427531
37	626.5	60.2	890	9	AL538796
38	626	60.1	1078	13	BM541636
39	615	59.1	939	9	AL554686
40	614	59.0	898	14	BQ733563
41	613.5	58.9	863	14	BQ894646
42	612	58.8	906	9	AL544313
43	606	58.2	599	13	BQ064020
44	604	58.0	664	13	BJ061844
45	601	57.7	596	13	BJ058360

ALIGNMENTS

RESULT 1
BQ720798

LOCUS

DEFINITION

AGENCOURT_8230839 Lupski_sympathetic_trunk Homo sapiens cdna clone

IMAGE:619T156 5', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

925 bp mRNA linear EST 16-JUL-2002

GI:21859695

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 925)

NIH-MGC http://mgs.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Qy 21 GluValPheThrAsnGlnGluValLysGluLysPheGluGlyLeuPheArgThrTyAsp 40
 Db 252 GAGGTCCTTTACCAATCAGAGGTTAAGAAATTTGAGGACTGTTTCGGACTTATGAT 311

Qy 41 GluCysValThrPheGlnLeuPheLysSerPheArgArgValArgGlyLeuAsnPheSerHis 60
 Db 312 GACTGTGTGACGTTCCAGCTATTTAAGAGTTTCAGACGTGTCGGTATATAAATTCAGCAAT 371

Qy 61 ProllysSerAlaAlaArgAlaArgGluGluLeuHisGluThrGlnPheArgGlyLysLys 80
 Db 372 CCTAAATCTGCAGCCCGAGCTAGGATAGCTTCATGAACCCCAATTCAGAGGGGAAAAA 431

Qy 81 LeuLysLeuTyPheAlaGlnValGlnThrProGluThrAspGlyAspLysLeuHisLeu 100
 Db 432 TTAAGCTCTACTTTGCACAGGTTCCAGACTCCAGAGACAGATGAGACAACTGCACCTG 491

Qy 101 AlaProGlnProAlaLysGlnPheLeuLysSerProProSerSerProProValGly 120
 Db 492 GCTCACCCAGCTGCCAACAGTTTCTCATCTGCCCCCTTCTCCACCTGTTGGC 551

Qy 121 TrpLysProLysSerAspAlaThrProValLeuAsnTyPheLeuLeuTyPheAlaValAla 140
 Db 552 TGGCAGCCCATCAACGATGCCACCGAGTCTCACTATGACCTCTATGCTGTGGCC 611

Qy 141 LysLeuGlyProGlyGluLysTyGluLeuHisAlaGlyThrGluSerThrProSerVal 160
 Db 612 AACTAGACACGAGAGAGATGAGCTCCATGACGGACTGATGCCACCCCAAGTGTC 671

Qy 161 ValValHisValCysAspSerAspMetGluGluGluGluAspProLysTyPheSerProLys 180
 Db 672 GTCTGCAGCTGTCGACAGTGACATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 731

Qy 181 ProllysLeuGlnThrArgArgProGlyLeu-ProProSerValSerAsn 197
 Db 732 CCAAAATATCATCCAACTCGCGCTGCTGGCTGCCACCTCGGTGCCAAC 783

RESULT 3
 AL532776 903 bp mRNA linear EST 13-FEB-2001
 LOCUS LTI_FL015_Brn1 Homo sapiens cDNA clone CS0DN001Y008 5
 DEFINITION Prime, mRNA sequence.
 ACCESSION AL532776
 VERSION AL532776.1 GI:12796269
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 903)
 AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES

Source

1. .903
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS0DN001Y008"
 /clone_lib="LTI_FL015_Brn1"
 /sex="male"
 /tissue_type="Adult brain"
 /note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with NotI and cloned into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library was constructed by Life Technologies.
 Contact: Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax: (1) 301 610 8371 Email: fliang@lifetech.com URL:
 B1545609 847 bp mRNA linear EST 05-SEP-2001
 603187809P1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:5259147 5',
 mRNA sequence.
 B1545609
 ACCESSION B1545609
 VERSION B1545609.1 GI:15432921
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 847)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

BASE COUNT 222 a 249 c 206 g 221 t 5 others
 ORIGIN

Alignment Scores:
 Pred. No.: 1.65e-98 Length: 903
 Score: 994.00 Matches: 189
 Percent Similarity: 98.48% Conservative: 6
 Best Local Similarity: 95.45% Mismatches: 2
 Query Match: 95.49% Indels: 1
 DB: Gaps: 0

US-09-782-953-9 (1-197) x AL532776 (1-903)

Qy 1 MetProAlaProSerMetAspCysAspValSerThrLeuValAlaCysValValAspVal 20
 Db 171 ATGCCAGCCCTAGCATGAGCTGTGATGTTTCCACTCTGGTGGCTGTGTGGGATGTC 230

Qy 21 GluValPheThrAsnGlnGluValLysGluLysPheGluGlyLeuPheArgThrTyAsp 40
 Db 231 GAGGTCTTTACCAATCAGAGGTTAAGAAATTTGAGGACTGTTTCGGACTTATGAT 290

Qy 41 GluCysValThrPheGlnLeuPheLysSerPheArgArgValArgLysLeuHisSerHis 60
 Db 291 GACTGTGTGACGTTCCAGCTATTTAAGAGTTTCAGACGTGTCGGTATATAAATTCAGCAAT 350

Qy 61 ProllysSerAlaAlaArgAlaArgGluGluLeuHisGluThrGlnPheArgGlyLysLys 80
 Db 351 CCTAAATCTGCAGCCCGAGTAAAGATAGAGTTCATGAACCCCAATTCARAGGGGAAAAA 410

Qy 81 LeuLysLeuTyPheAlaGlnValGlnThrProGluThrAspGlyAspLysLeuHisLeu 100
 Db 411 TTAAGCTCTACTTTGCACAGGTTCCAGCTCCAGAGACAGATGAGACAACTGCACCTG 470

Qy 101 AlaProGlnProAlaLysGlnPheLeuLysSerProProSerSerProProValGly 120
 Db 471 GCTCCACCCAGACCTGCCAACAGTTTCTCATCTGCCCCCTTCTCCACCTGTTGG 530

Qy 120 TrpLysProLysSerAspAlaThrProValLeuAsnTyPheLeuLeuTyPheAlaValAla 140
 Db 531 CTGGCAGCCCATCAACGATGCCACCGAGTCTCACTATGACCTCTCTATGCTGTGGC 590

Qy 140 aLysLeuGlyProGlyGluLysTyPheGluLeuHisAlaGlyThrGluSerThrProSerVa 160
 Db 591 CAACTAGGACACGAGAGAGATGAGTCCATCGAGGACTGAGTCCACCCCAAGTGT 650

Qy 160 lValValHisValCysAspSerAspMetGluGluGluGluAspProLysTyPheSerProLy 180
 Db 651 COTCTGTGACGTTGTCGACAGTGCATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 710

Qy 180 sProllysLeuGlnThrArgArgProGlyLeuProProSerValSerAsn 197
 Db 711 GCCAAATATCATCCAACTCGCGCTGCTGGCTGCCACCTCGGTGCCAAC 762

RESULT 4

B1545609

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT


```

Db 303 TATGACCTCCTATGCTGTGGCCAAACTAGGACCAGGAGAAAGTAGTGTGCTCCATGCA 362
Qy 153 GlyThrGluSerThrProSerValValHisValCysAspSerAspMetGluGluGlu 172
Db 363 GGGACTGAGTCCACCCCAAGTGTGCTGTGACGTGTGCGACAGTGTGACATAGAGAGAA 422
Qy 173 GluAspProLysThrSerProLysProLysProLysLeuLeuLeuLeuLeuLeuLeu 192
Db 423 GAGACCCAAAGACTTCCCAAGCCAAATCATCCAAACTCGGCGCTGTGGCTGCCA 482
Qy 193 ProSerValSerAen 197
Db 483 CCCTCCGTGTCCAC 497

RESULT 6
BG709000
LOCUS
DEFINITION
602675149F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4797589 5',
mRNA sequence.
ACCESSION
BG709000
VERSION
EST.
KEYWORDS
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 758)
NIH-MGC http://mgc.nci.nih.gov/.
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10684 row: a column: 14
High quality sequence stop: 746.
Location/Qualifiers
1..758
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4797589"
/clone_lib="NIH MGC 96"
/tissue_type="hypothalamus"
/lab_host="DH108"
/note="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTATTATTATTATTVN-3',
size-selected for average insert size 2.3 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 184 a 206 c 199 g 169 t
ORIGIN
1..758
x

```

US-09-782-953-9 (1-197) x BG709000 (1-758)

Qy 1 MetProAlaProSerMetAspCysAspValSerThrLeuValAlaCysValValAspVal 20

```

Db 236 ATGCCAGCCCTAGCTAGCTGCTGATGTTTCCACTCTGTTGCCTGTGTGGTGAATGC 295
Qy 21 GluValPheThrAsnGlnGluValLysGluLysPheGluGluPheArgThrThrAsp 40
Db 296 GAGGTCTTTTACCCTAATCAGGAGGTTAAAGAAAATTTTGGGAGCTGTTTCGGACTTATGAT 355
Qy 41 GluCysValThrPheGlnLeuPheLysSerPheArgValArgLysLeuPheArg-GlyLys 60
Db 356 GACTGTGTGACGTTTCCAGCTATTAGAGTTTCAGACGTTGTCGTATTAACATTACCAAT 415
Qy 61 ProLysSerAlaAlaArgAlaArgLysGluLeuHisGluThrGlnPheArg-GlyLys 80
Db 416 CCTAAATCTGCAGCCGAGCTAGGATAGAGTTCATGAACCAATTCAGAGGGACAAA 475
Qy 80 stLeuLysLeuThrPheAlaGlnValGlnThrProGluThrAspGlyAspLysLeuHis 100
Db 476 ATTAAAGCTCTACTTTGTCAGAGTTTCAGACTCCAGAGACAGATGGAGACAAATTCGACTT 535
Qy 100 uAlaProGlnProAlaLysGlnPheLeuLeuSerProProSerSerProProValG 120
Db 536 GGCTCCACCCAGCTGCCAAACAGTTTCTCATCTCGCCCTTCTCTCCCTCTCTCTCTCT 595
Qy 120 YTrpLysProLysSerAspAlaThrProValLeuAsnThrAspLeuLeuThrAlaVal 140
Db 596 CTGGCAGCCCATCAACGATGCCACGCTCTCAACTATGACCTCTCTCTCTCTCTCTCTCT 655
Qy 140 aLysLeuGlyProGlyGluLysThrGluLeuHisAlaGlyThrGluSerThrProSerVa 160
Db 656 CAACTAGGACCCAGGAGAGATGATGAGTCCATGCGGAGCTGAGTCTCTACCAAGTGT 715
Qy 160 lValValHisValCysAspSerAspMetGlu 170
Db 716 CTGCTGTCAGCTGTGCGACAGTGCATAGAG 746

RESULT 7
BF443461
LOCUS
DEFINITION
260951 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence.
ACCESSION
BF443461
VERSION
BF443461.1 GI:11503553
KEYWORDS
EST.
SOURCE
pig.
ORGANISM
Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE
1 (bases 1 to 495)
Fahrenkrug, S.C., Freking, B.A., Rohrer, G.A., Smith, T.P.L., Casas, E.,
Stone, R.T., Heaton, M.P., Grosse, W.M., Bennett, G.A., Laegreid, W.W.
and Keele, J.W.
Design and use of two pooled tissue normalized cDNA libraries for
EST discovery in swine
Unpublished (2000)
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCCTCAGTCACGACG
Plate: 93 row: G column: 19
Seq primer: ATTTAGGTGACATATAG.
Location/Qualifiers
1..495
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC 2P1G"
/tissue_type="pooled"
FEATURES
source

```

/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from testis, ovary,
endometrium, hypothalamus, pituitary, and placenta."
BASE COUNT 134 a 145 c 122 g 94 t
ORIGIN

Alignment Scores:
Pred. No.: 3,766-78 Length: 495
Score: 805.00 Matches: 158
Percent Similarity: 98.79% Conservative: 5
Best Local Similarity: 95.76% Mismatches: 1
Query Match: 77.33% Indels: 2
DB: 12 Gaps: 0

US-09-782-953-9 (1-197) x BF443461 (1-495)

OY 20 ValGluValPheThrMetGlnGluValValGluValPheGluGluValPheArgThrTyr 39
DB 2 GTGGAGGCTTTTACCAACGAGGTTAAGA-AATTGAGGGCTATTCCGACTTAT 60
OY 40 AAGGluCysValThrPheGlnLeuPheLysSerPheArgValArgLLeuPheSer 59
DB 61 GACGAGTGTGTGACGTTCCAGCTGTTTAAAGACTTCAGACGTCTCCGAATTAATTCAGC 120
OY 60 HisProLysSerAlaAlaArgAlaArgLLeuLeuHisGluThrGlnPheArg-GlyLys 79
DB 121 AATCCCAATCTGCACCCGAGCCAGATAGACTCATGAAACCCAGTTCAAGGGGGA 180
OY 79 GlyValLeuValPheThrPheAlaGlnValGlnThrProGluThrAspGlyAspLysLeuHis 99
DB 181 AAAATTAAACTCTTACTTGTGCACAGGTTCAGACTCCAGACAGACGAGGAGACAACTGCA 240
OY 99 GluAlaProGlnProGlnProAlaLysGlnPheLeuLysSerProPheSerProProVa 119
DB 241 TTGGGCCCAACCAAGCTCCCAACATTCCTCATCTCACCCCGCTCTCTCCCT 300
OY 119 GGTTPLyProLysSerAspAlaThrProValLeuAsnTyrAspLeuLysAlaVal 139
DB 301 CGGCTGGCAGCCCATCAGCATGCAAGCGCGTCTTAACTAGACCTCTATGCCGT 360
OY 139 AlaValLeuGlyProGlyGlyLysTyrGluLeuHisAlaGlyThrGlnSerThrProSe 159
DB 361 GGCCAAACTAGAGCAGAGAGAAAGTACGAGCTCCACGCGAGGAGCGAGTCCACCCGAG 420
OY 159 ValValValHisValCysAspSerAspMetGluGluGluGluAspProLysThrSerPr 179
DB 421 TGTCTGTGACAGCTGTGTGACATGATCAGAGAGAAAGATCCAAAGACGTCCCC 480
OY 179 OLYsProLYsIle 183
DB 481 CAAGCCAAAAATC 493
RESULT 8
LOCUS BG721994 742 bp mRNA linear EST 08-MAY-2001
DEFINITION 602698760F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4830759 5',
ACCESION BG721994
VERSION BG721994.1 GI:14001181
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 742)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bcr-remail.nih.gov
Tissue Procurement: Miklos Palcovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki

Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LRAM10752 row: g column: 16
High quality sequence stop: 738.
Location/Qualifiers

FEATURES

source

1.742
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4830759"
/clone_id="NIH_MGC_97"
/lab_host="DH10B"
/note="Organ: testis; Vector: pBluescript (modified
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (GTCGAG
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size selected for average insert size 2.2 kb and
normalized to 10^5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT 180 a 202 c 193 g 167 t
ORIGIN

Alignment Scores:

Pred. No.: 1,056-77 Length: 742
Score: 803.00 Matches: 164
Percent Similarity: 98.26% Conservative: 5
Best Local Similarity: 95.35% Mismatches: 3
Query Match: 77.14% Indels: 3
DB: 12 Gaps: 0

US-09-782-953-9 (1-197) x BG721994 (1-742)

OY 1 MetProAlaProSerMetCysAspValSerThrLeuValAlaCysValValAspVal 20
DB 228 ATGCCAGCCCTTACAGATGATGATGATTTCCACTGTGCTGCTGTGTGATGATCTC 287
OY 21 GluValPheThrAsnGluValValGluLysPheGluGlyLysPheArgThrTyrAsp 40
DB 288 GAGGCTTTTCAATCAGAGGTTAAGAAATTTGAGAGCTGTTCCGACTTATAT 347
OY 41 GluCysValThrPheGlnLeuPheLysSerPheArgValArgLLeuPheSerHis 60
DB 348 GACTGTGTGACGTTCCAGCTATTTAAAGTTTCAGACGTGCCGATTAATCTCAGCAT 407
OY 61 ProLysSerAlaAlaArgAlaArgLLeuLeuHisGluThrGlnPheArgGlyLysLys 80
DB 408 CCTTAATCTGCAGCCCGAGCTAGTAGAGCTTCATGAAACCAATTCAGAGGAAAAA 467
OY 81 LeuValLeuTyrPheAlaGlnValGlnThrProGluThrAspGlyAspLysLeuHisLeu 100
DB 468 TTTAAAGCTTACTTGTGCACAGGTTCAGACTCCAGACAGATGAGACAACTGCACCTTG 527
OY 101 AlaProGlnProAlaLysGlnPheLeuLysSerProPheSerProProVaGly 120
DB 528 GCTCCACCCCAAGCTGCAAGTTTCTCATCTGCCGCCCTCTCCACCTGTGGC 587
OY 121 TTPLyProLysSerAspAlaThrProValLeuAsnTyrAspLeuLysValAla 140
DB 588 TGGCAGCCCATCAGATGCCAGCGCAGTCCCA-TATGACCTCTCTATGCTGTGGC- 645
OY 141 LysLeuGlyProGlyGlyLysTyrGluLeuHisAlaGlyThrGlnSerThrProSerVal 160
DB 646 AACTAGAGCAGAGAGAGATAGCTCCATGAGGAGACTGATCCAC-CCAAAGTGTTC 704
OY 161 ValValValHisValCysAspSerAspMetGluGluGlu 172
DB 705 GTCTGTGACGTTGTGACAGTACATAGAGAGAA 740

RESULT 9

BQ260356

LOCUS

DEFINITION

BQ260356 546 bp mRNA linear EST 06-MAY-2002
faa10f12.y1 zebrafish fin day3 regeneration Danio rerio cDNA clone
5909542 5' similar to SW:ZAK4_HUMAN Q14206 ZAK1-4 PROTEIN. ;, mRNA
sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

BQ260356.1

GI:20461126

zebrafish.

EST.

Danio rerio

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes

; Cyprinidae; Danio.

1 (bases 1 to 546)

Clark, M., Johnson, S.L., Lehrach, H., Lee, R., Li, F., Marra, M., Eddy

, S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood

, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B.,

Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E.,

Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R.

and Wilson, R.

WashU Zebrafish EST Project 1998

Unpublished (1998)

Contact: Stephen L. Johnson

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: zbrfish@wustl.edu

cDNA Library Preparation: Raymond Lee. cDNA Library Arrayed by:

Matthew Clark. DNA Sequencing by: Washington University Genome

Sequencing Center Clone distribution: Genome Systems, St. Louis,

Missouri (web address: www.genomesystems.com) (email contact:

info@genomesystems.com) and Research Genetics, Huntsville, Alabama

(web address: www.resgen.com) (email contact: info@resgen.com) and

ReSourceZentrumPrimarDatenbank, Berlin, Germany (web address:

www.rzpd.de)

Seq primer: T3 ET from Amersham

High quality sequence stop: 395.

Location/Qualifiers

1..546

/organism="Danio rerio"

/db_xref="taxon:7955"

/clone="5909542"

/clone_lib="zebrafish fin day3 regeneration"

/sex="mixed male and female"

/tissue type="3 day fin regenerates"

/lab host="E. coli XL0LR"

/note="Vector: pBK-CMV; Site 1: EcoRI; Site 2: XhoI; 1st

strand cDNA primed with (GA)10ACTAGTCTCGAG(T)18, followed

by second strand synthesis, and ligated to 5' adapter (5'

)-aattcgacag-3', 3'-gccgtgctc-5'. cDNA was cloned

directionally (EcoRI/XhoI) into Stratagene zap express

lambda phage arms. Mass in vivo excision done to obtain

inserts in pBK-CMV phagemid."

144 a 151 c 139 g 112 t

BASE COUNT

ORIGIN

Alignment Scores:

Pred. No.: 3,1e-76

Score: 788.00

Length: 546

Matches: 148

Percent Similarity: 89.56%

Conservative: 15

Best Local Similarity: 81.32%

Mismatch: 19

Query Match: 75.70%

Indels: 0

DB: 14

Gaps: 0

US-09-782-953-9 (1-197) x BQ260356 (1-546)

QY 13 LeuValAlaCysValValAspValGluValPheThrAsnGlnGluValLysGluLysPhe 32

Db 1 CTATCGCCTGTCGGTGGACGTGGAGGTGTTCCACGATCGAGAGGTCAAGGAATGTTT 60

QY 33 GluGlyLeuPheArgThrTyraSpGluCyvalThrPheGlnLeuPheLysSerPheArg 52

Db 61 GAAGCGCTCTTCTCGCTACGATGAGGTGTGACGTTTCAGAGCTTTCAGA 120
QY 53 ArgValArgIleAsnPheserHisProLysSerAlaAlaArgAlaArgIleGluLeuHis 72
Db 121 CGTGTACGAAATCAATTTAGTAACTTAAAGCAGCAGCCAGGCAAGATAGAGCTGCAC 180
QY 73 GluThrGlnPheArgGlyLysLysLeuLysLeuTyraPheAlaGlnValGlnThrProGlu 92
Db 181 GAGACGCTTTTCAGAGGGAAGAACTCAAACTTTTACTTTTCCAGGTGAGAAACCCAGTA 240
QY 93 ThrAspGlyAspLysLeuHisLeuAlaProProGlnProAlaLysGlnPheLeuLysSer 112
Db 241 TCTGAAGGAGACAATTTGACCTGCTCTCCACAGCCTTCCAAACAGTTCCTCATCTCT 300
QY 113 ProProSerSerProProValGlyTrpLysProLysSerAspAlaThrProValLeuAsn 132
Db 301 CCCCTGGCTCACTCCAGTCGTTGGCAACAGATTGATGAGCCACACCGGTTATTAAAC 360
QY 133 TyrAspLeuLeuTyraAlaValAlaLysLeuGlyProGlyGluLysTyraGluLeuHisAla 152
Db 361 TACGACCTGCTGTATGACGTTGCCAAGCTTGGCCAGGTGAGAAGTTCGAGCTCCATCG 420
QY 153 GlyThrGluSerThrProSerValValHisValCysAspSerAspMetGluGluGlu 172
Db 421 GGCACAGAGTGTACGCCAGTGTGTTGCTCCACGTTCTGTGACAGGACACGATGAGGAG 480
QY 173 GluAspProLysThrSerProLysProLysIleIleGlnThrArgArgProGlyLeuPro 192
Db 481 GAGGACCAACAAAGAACACCTTAACCAAAATATCCAGACCCGACCTCCAGACCTGCCA 540
QY 193 ProSer 194
Db 541 GTCTCA 546

RESULT 10

B1476888

LOCUS

DEFINITION

B1476888

617 bp mRNA linear EST 27-AUG-2001

fq95g03.y3 zebrafish fin day3 regeneration Danio rerio cDNA clone

4926509 4926509 5' similar to SW:ZAK4_HUMAN Q14206 ZAK1-4 PROTEIN.

; mRNA sequence.

ACCESSION

B1476888

VERSION

B1476888.1

GI:15310227

KEYWORDS

EST.

SOURCE

zebrafish.

ORGANISM

Danio rerio

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes

; Cyprinidae; Danio.

1 (bases 1 to 617)

Clark, M., Johnson, S.L., Lehrach, H., Lee, R., Li, F., Marra, M., Eddy

, S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood

, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B.,

Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E.,

Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R.

and Wilson, R.

WashU Zebrafish EST Project 1998

Unpublished (1998)

Contact: Stephen L. Johnson

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: zbrfish@wustl.edu

cDNA Library Preparation: Raymond Lee. cDNA Library Arrayed by:

Matthew Clark. DNA Sequencing by: Washington University Genome

Sequencing Center Clone distribution: Genome Systems, St. Louis,

Missouri (web address: www.genomesystems.com) (email contact:

info@genomesystems.com) and Research Genetics, Huntsville, Alabama

(web address: www.resgen.com) (email contact: info@resgen.com) and

ReSourceZentrumPrimarDatenbank, Berlin, Germany (web address:

www.rzpd.de)

High quality sequence stop: 467.

to Rot = 20.0 and subtraction to Rot = 479.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATTCTGAGTTAATAATTAATCCCTCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda PLC I."

BASE COUNT 149 a 183 c 158 g 150 t
ORIGIN

Alignment Scores:
Pred. No.: 1,3e-74 Length: 640
Score: 774.00 Matches: 147
Percent Similarity: 99.32% Conservatives: 0
Best Local Similarity: 99.32% Mismatches: 1
Query Match: 74.35% Indels: 0
DB: 10 Gaps: 0

US-09-782-953-9 (1-197) x BB550331 (1-640)

QY 1 MetProAlaProSerMetAspCysAspValSerThrLeuValAlaCysValValAspVal 20
DB 197 ATGCCAGCCCTAGCATGGACTGTGATTTTCCACTCTGTCGCCCTGTGTGGTGTG 256
QY 21 GluValPheThrAsnGlnGluValLysGluLysPheGluLysPheArgThrTyRAsp 40
DB 257 GAGTCTTTACCATTCAGAGGTTAAGGAAATTCGAGGACTGTTCGGACCTATGAT 316
QY 41 GluCysValThrPheGlnLeuPheLysSerPheArgValArgIleAsnPheSerHis 60
DB 317 GAATGTGTGAGGTTCCAGCTGTTTAAGAGTTTCCGAGGTTTCGATAAATTCAGCCAT 376
QY 61 ProLysSerAlaAlaArgAlaArgIleGluLeuHisGluThrGlnPheArgGlyLysLys 80
DB 377 CCCAAATCTGCAGCCCGTCCCGGATAGAGCTTCATGAGACTCAGTTCAGAGGGAAG 436
QY 81 LeuLysLeuTyR-PheAlaGlnValGlnThrProGluThrAspGlyAspLysLeuHisLeu 100
DB 437 CTAAACTCTACTTCCGCCAGGTCAGAGCCAGAGACAGATGGAGACAACTGCATTTG 496
QY 101 AlaProGlnProAlaLysGlnPheLeuIleSerProProSerSerProProValGly 120
DB 497 GCACCTCCAGCGCTGCCAACAAGTTCCTCACTCACCCCTTCATCTCTCTGTGGC 556
QY 121 TrpLysProIleSerAspAlaThrProValLeuAsnTyRAspLeuLeuTyRAlaValAla 140
DB 557 TGGAGCCTATCAGCGATGCCACACAGTCTCAACTATGACCTTCTTATGCTGTGGCC 616
QY 141 LysLeuGlyProGlyGluLysTyR 148
DB 617 AAACCTAGGACCGAGGAGAAATAT 640

RESULT 12
BF534085
LOCUS 602075073F1 NCI_CGAP_Li9 Mus musculus cDNA clone IMAGE:4211994 5',
DEFINITION mRNA sequence.

ACCESSION BF534085
VERSION BF534085.1 GI:11621448

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS 1 (bases 1 to 1007)

TITLE NIH-MGC http://mgi.nci.nih.gov/

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing By: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Plate: L1AM9782 row: a column: 19

High quality sequence stop: 742.

FEATURES

source

1..1007
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4211994"
/clone_lib="NCI CGAP_Li9"
/lab_host="DH10B (TI phase-resistant)"
/note="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.9 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT 249 a 283 c 255 g 220 t

ORIGIN

Alignment Scores:

Pred. No.: 1,16e-72 Length: 1007
Score: 758.50 Matches: 159
Percent Similarity: 86.46% Conservatives: 7
Best Local Similarity: 82.81% Mismatches: 20
Query Match: 72.86% Indels: 6
DB: 12 Gaps: 1

US-09-782-953-9 (1-197) x BF534085 (1-1007)

QY 11 SerThrLeuValAlaCysValValAspValGluValPheThrAsnGlnGluValLysGlu 30
DB 264 AACTCATTTGTTGCTGCAATGTTCCACGAGTCTGTGTTTGAAGAAGAGGAGGCAAGAA 323
QY 31 LysPheGluGlyLeuPheArgThrTyRAspGluCysValThrPheGlnLeuPheLysSer 50
DB 324 AAATTCGAGGGGACTGTTCCGGACCTATGATGAATGTGTGACGTTCCAGCTGTTAAGAGT 383
QY 51 PheArgAlaGValArgIleAsnPheSerHisProLysSerAlaAlaArgAlaArgIleGlu 70
DB 384 TTCCAGCGGTTTCGATAAATTTTCAGCCATCCCAATCTGCAGCCCGTCCCGGATAGAG 443
QY 71 LeuHisGluThrGlnPheArgGlyLysLysLeuLysLeuTyR-PheAlaGlnValGlnThr 90
DB 444 CTTCATGAGACTCAGTTCAGAGGGAAGAAGCTAAACTCTACTTCGCCAGGTCCAGACC 503
QY 91 ProGluThrAspGlyAspLysLeuHisLeuAlaProGlnProAlaLysGlnPheLeu 110
DB 504 CCAGAGACAGATGGAGACAACTGCAITTTGGCACCCTCCACAGCCTGCCAAACAGTTCCTC 563
QY 111 IleSerProProSerSerProProValGlyTyrLysProIleSerAspAlaThrProVal 130
DB 564 ATCTCACCCCTTCATCTCTCTCTCTGCTGGAGAGCCCTATCAGCGATGCCACACGATC 623
QY 131 LeuAsnTyRAspLeuLeuTyRAlaValAla-LysLeuGlyProGly-GluLysTyRGlul 150
DB 624 CTCAACTATGACTCTTTATGCTGTGGCCAAACTAGGACCCAGGAGCAGAGTATGAGC 683
QY 150 euHisAlaGlyThrGluSerThrProSerValValHisValCysAsp--SerAspMe 169
DB 684 TGCACGCTGGAACTAGTCTACACCGAGCGTGTGTGTGCTGTGTGTGACAGCGACACT 743
QY 169 tGluGluGluGluAspProLysThrSerProLysProLysIleIleGlnThrArgArgPr 189
DB 744 GCGCGCGGAAGAACCCCAAGAAATT---CCCAAGCCAAAATCATTTCCAGACCGGAGTCC 800
QY 189 oGlyLeu-ProProSerValSerAsn 197
DB 801 TGGCTGGTCAACCATCGGGTCCAAT 826
RESULT 13
AI796134/c
LOCUS AI796134 557 bp mRNA linear EST 06-JUL-1999

DEFINITION	wh42909.x1 NCI CGAP Kid11 Homo sapiens cDNA clone IMAGE:2383456 3
ACCESSION	A1796134
VERSION	A1796134.1
KEYWORDS	EST.
ORGANISM	Homo sapiens
SOURCE	human.
COMMENT	Euryarchaea; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 557) NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap . National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index Unpublished (1997) Contact: Robert Strausberg, Ph.D. Email: cgapb@remail.nih.gov Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D. CGNA Library Preparation: M. Bento Soares, Ph.D. CGNA Library Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/bbrp/image/image.html Seq primer: -40up from Gibco High quality sequence group: 455. Location/Qualifiers
FEATURES	1..557
SOURCE	/organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:2383456" /clone_lib="NCI-CGAP_Kid11" /lab_host="DH10B" /note="Organ: kidney; Vector: pTR73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not 1; Site 2: Eco RI; plasmid DNA from the normalized library NCI CGAP Kid3 was prepared, and as circles were used in vitro. Following HAT hybridization reaction, this DNA was used as tracer in a subtractive library construction reaction. The driver was PCR-amplified cDNA from a pool of 5,000 clones made from the same library (cloneids 1322376-1323911, 1456007-1456775, and 1500552-1502855). Subtraction by Bento Soares and M. Fatima Bonaldo."
BASE COUNT	125 a 107 c 162 g 162 t 1 others
ORIGIN	
Alignment Scores:	
Pred. No.:	7,61e-72 Length: 557
Score:	748.00 Matches: 141
Percent Similarity:	97.35% Conservative: 6
Best Local Similarity:	93.36% Mismatches: 4
Query Match:	71.85% Indels: 0
DB:	Gaps: 0
US-09-782-953-9 (1-197) x A1796134 (1-557)	
Oy	46 GInleuphelySeSerphearGArGValAArgileAsnPheserhiabProlySeraIaala 65
Db	557 CAGCATTATTAAGAGTTTCAACAGTCGCGATTAACCTTCAGCAATCTTAATCTGCAGCC 498
Oy	66 ArGAlaargIleGleuLunhIegIuthrGInpnearGlyLylyleyleuLyLeuLytrPhe 85
Db	497 CGAGCTAGAGATAGAGCTTCATGAAACCCCAATTCAGAGGGGAAAAAATTAAGGCTCTACTT 438
Oy	86 AlAGInvalAGlThrProGluThrArgGlyYaaRlyleuHiIleuNlaabProGInpRo 105
Db	437 GCACAGGTTCAAGCTCCAGAGACAGATGAGAAATGCACTTGGCTCCACCCAGCCCT 378
Oy	106 AlAtyGInPhleuIleSeSerProPSeSerSerProProValGlyYrPlyPProIleSer 125
Db	377 GCCAAACAGTTTTCATCTCGCCCCCTTCTCTCCCAACCTGTTGGCTGGCAGCCCAATCAAC 318
Oy	126 AapAlaThrProValIleuNanTyraBpLeuLeuTyraIaIaIaLyLeuGlyProGly 145

Db	317	GATGCACGCCAGTCTCTCACTATGACCTCTCTATGCTGTGGCCAAACTGAGACAGGA	258
Qy	146	GlutylrTgTlueHhAlglYThrGlueThrProsevalValIvalhIvalYs	165
Db	257	AAAAAGTATGAGCTTCATGACGAGGAGCTAGTCCACCCCAAGATGTCGTGTCACCTGTGC	198
Qy	166	AspserAspMetGluGluGluAspProIyThrsrProIyProIyIleIleGln	185
Db	197	GACGTGACATAGAGGAAAGAGAGGACCAAGATCTCCCAAGGCCAAATAATCATCCAA	138
Qy	186	ThraArgProGlyLeuProProsevalser	196
Db	137	ACTCGGGTCTCTGCGCTCCACCTCCCTGTCTC	105
RESULT 14			
LOCUS	BG701025	626 bp	mRNA linear EST 07-MAY-2001
DEFINITION	602682030P1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:4814743 5',		
ACCESSION	BG701025		
VERSION	BG701025.1	GI:13970954	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
TITLE	NIH-MGC http://mgc.nci.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapbs-rt@mail.nih.gov		
	Tissue Procurement: Miklos Palokovitz, M.D., Ph.D.		
	cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shizaki		
	Toshiyuki and Piero Carninci (RIKEN)		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LNLN at:		
	http://image.lnl.gov		
	Plate: LRAM10710 row: 1 column: 08		
	High quality sequence stop: 626.		
FEATURES			
source	Location/Qualifiers		
	1..626		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:4814743"		
	/clone_lib="NIH_MGC_95"		
	/tissue_type="hippocampus"		
	/lab_host="DH10B"		
	/note="Organ: brain; Vector: pBluescriptpr (modified		
	pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (Grcgag		
) ; Oligo-dT primed using primer 5'-TTTTTTTTTTTTTNN-3',		
	size-selected for average insert size 2.5 kb and		
	normalized to R0T 5. This is a primary library enriched		
	for full-length clones and constructed using the		
	Cap-trapper method (Carninci, in preparation). Library		
	constructed by M. Brownstein (NIH/NHGRI, National		
	Institutes of Health). Note: this is a NIH_MGC Library."		
BASE COUNT	150 a 177 c 154 g 145 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	1,37e-69	Length:	626
Score:	728.00	Matches:	138
Percent Similarity:	99.30%	Conservative:	4
Best Local Similarity:	96.50%	Mismatches:	1
Query Match:	69.93%	Indels:	0
DB:	12	Gaps:	0

```

QY 1 MetProAlaProSerMetAspCysAspValSerThrLeuValAlaCysValValAspVal 20
Db 196 ATGCCAGCCCTAGCATGGACTGTGATGTTTCCACTCTGGTTGGCTGTGTGGATGTC 255
QY 21 GluValPheThrAsnGlnGluValLysGluLysPheGluGlyLeuPheArgThrTyAsp 40
Db 256 GAGGTCTTTACCAATCAGGAGGTTAAGGAAAATTTGAGGACGTGTTCCGCACTTATGAT 315
QY 41 GluCysValThrPheGlnLeuPheLysSerPheArgValArgIleAsnPheSerHis 60
Db 316 GACTGTGTGAGCTTCCAGCTATTTAAGAGTTTCAGACGTGTCCGTATATAAATTCAGCAAT 375
QY 61 ProLysSerAlaAlaArgAlaArgIleGluLeuHisGluThrGlnPheArgGlyLysLys 80
Db 376 CCTAAATCTGAGCCCGAGCTAGGATAGAGCTTCATGGAACCCCAATTCAGAGGGGAAAAA 435
QY 81 LeuLysLeuTyPheAlaGlnValGlnThrProGluThrAspGlyAspLysLeuHisLeu 100
Db 436 TTAAGCTCTACTTTGCACAGTTTCAGACTCCAGACAGATGGAGACAACTGCATTTG 495
QY 101 AlaProGlnProAlaLysGlnPheLeuLysSerProProSerSerProProValGly 120
Db 496 GCTCACCCAGCTGCGCAACAGTTTCTCATCTCGCCCTTCTCTCCACCTGTTGGC 555
QY 121 TrpLysProLysSerAspAlaThrProValLeuAsnTyPheLeuLeuTyPheAlaValAla 140
Db 556 TGGCAGCCCATCAACGATGCGACCCAGTCTCACTATGACCTCTCTATGCTGTGGCC 615
QY 141 LysLeuGly 143
Db 616 AAAC TAGGA 624

```

```

RESULT 15
AUI69846 747 bp mRNA linear EST 29-JAN-2001
LOCUS AUI69846 Ol-br-ad cDNA Oryzias latipes cDNA clone br5301, mRNA
DEFINITION

```

```

ACCESSION AUI69846
VERSION AUI69846
KEYWORDS AUI69846.1 GI:12591915
SOURCE EST.
ORGANISM Japanese medaka.
Oryzias latipes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
1 (bases 1 to 747)
Mita, K., Iehikawa, Y. and Yamauchi, M.
Establishment of cDNA database of medaka, Oryzias latipes
Unpublished (2001)
Contact: Mita K
Genome Research Group
National Institute of Radiological Sciences
Anagawa 4-9-1, Inage, Chiba 263-8555, Japan
Email: kmita@nirs.go.jp
method: uni-directional sequence direction: sequenced from T3 primer
(5' -> 3').

```

```

FEATURES
source Location/Qualifiers
1..747
/organism="Oryzias latipes"
/strain="HNT"
/db_xref="taxon:8090"
/clone="br5301"
/clone_lib="Ol-br-ad cDNA"
/sex="female/male mixed"
/tissue type="brain"
/dev_stage="adult"

```

```

BASE COUNT 193 a 210 c 191 g 153 t
ORIGIN

```

```

Alignment Scores: 2.85e-69 Length: 747
Pred. No.: 726.00 Matches: 138
Score:

```

```

Percent Similarity: 82.74% Conservative: 25
Best Local Similarity: 70.05% Mismatches: 32
Query Match: 69.74% Indels: 2
DB: 9 Gaps: 1

```

```

US-09-782-953-9 (1-197) x AUI69846 (1-747)

```

```

QY 1 MetProAlaProSerMetAspCysAspValSerThrLeuValAlaCysValValAspVal 20
Db 155 ATGCCATCGATAAATATGAACCTGGAGATGGTCAACCATCATGCCCTGCCAGTAGACACC 214
QY 21 GluValPheThrAsnGlnGluValLysGluLysPheGluGlyLeuPheArgThrTyAsp 40
Db 215 GAAGTGTTTACGTTCAGACAAAATGCGAGAGAGGTTTGAAGCCGTGTTAGGATTTATGAT 274
QY 41 GluCysValThrPheGlnLeuPheLysSerPheArgValArgIleAsnPheSerHis 60
Db 275 GAACACAACTTCCAGATGTTCAAAAGCTTCAGAAGATGCGGATCAACTTCAGCACC 334
QY 61 ProLysSerAlaAlaArgAlaArgIleGluLeuHisGluThrGlnPheArgGlyLysLys 80
Db 335 CCAGAGGCGAGCCCGCGCTCGCATCGAGTTGCACGAATCGGAGTTCAACGCGCAAGAAG 394
QY 81 LeuLysLeuTyPheAlaGlnValGlnThrProGluThrAspGlyAspLysLeuHisLeu 100
Db 395 CTCAAACTCTACTTTGCTCAAATCCAGAATGTTGACGAGGACATAGACAAGTCTTTACCTG 454
QY 101 AlaProGlnProAlaLysGlnPheLeuLysSerProProSerSerProProValGly 120
Db 455 GCGCCCTCCACAACCCCGTGAACACAGTTTCTGATCTCTCTCTCGCTCGCGCTGTGGGA 514
QY 121 TrpLysProLysSerAspAlaThrProValLeuAsnTyPheLeuLeuTyPheAlaValAla 140
Db 515 TGGAGCGCAGGTGAAGACGCCACCCGTCATCACTACGACCTGCTATGCTGTTGCC 574
QY 141 LysLeuGlyProGlyGlyLysTyPheGluLeuHisAlaGlyThrGluSerThrProSerVal 160
Db 575 AAGCTCGGACCTGGTGAAAGTATGAACCTCCACGCGGAAACAGAGTCCCACTCCAAGTGT 634
QY 161 ValValHisValCysAspSerAspMetGluGluGluAspProLysThrSerProLys 180
Db 635 GTCTGTCACGTTGCGATAGC-----GAGGCGAGGAGGACGAGCGGTGCGCCAAAG 688
QY 181 ProLysIleIleGlnThrArgArgProGlyLeuProProSerValSerAsn 197
Db 689 CAGCAGATGCTCCAGACCAAGCGGTCCCGATGCTCCCGACCGTCCCAAC 739

```

```

Search completed: December 14, 2002, 22:28:15
Job time : 1116.9 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 11:37:42 ; Search time 30.5632 Seconds
(without alignments)
858.887 Million cell updates/sec

Title: US-09-782-953-9

Perfect score: 1041

Sequence: 1 MPAPSMDCVSTLVACVDV.....SPKPKIQTTRPGLPPSVSN 197

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002:*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1041	100.0	197	23	AAE18913
2	980	94.1	192	23	AAE18916
3	905	86.9	234	22	AAE18917
4	905	86.9	255	23	AAE18918
5	658.5	63.3	241	23	AAE18917
6	641	61.6	197	23	AAE18914
7	633.5	60.9	212	23	AAE18918
8	632.5	60.8	198	23	AAE18912
9	622	59.8	197	23	AAE18915
10	613.5	58.9	198	23	AAE18911

11	586	56.3	171	20	AAW73898	Human Down's Syndr
12	424	38.8	292	22	ABB71467	Drosophila melanog
13	323	31.0	111	21	AAE18913	Human secreted pro
14	254.5	24.4	142	21	AAE18913	Lung cancer associ
15	219	21.0	58	22	ABB29495	Peptide #2146 enco
16	219	21.0	58	22	ABB34672	Peptide #2178 enco
17	219	21.0	58	22	ABB36877	Peptide #4383 enco
18	219	21.0	58	22	ABB20081	Protein #2080 enco
19	219	21.0	58	22	AAE18913	Human brain expres
20	219	21.0	58	22	AAE18913	Human brain expres
21	219	21.0	58	22	AAE18913	Human bone marrow
22	219	21.0	58	22	AAE18913	Peptide #2105 enco
23	219	21.0	58	22	AAE18913	Peptide #4293 enco
24	219	21.0	58	22	AAE18913	Peptide #2211 enco
25	219	21.0	58	22	AAE18913	Peptide #4402 enco
26	219	21.0	58	22	AAE18913	Peptide #2088 enco
27	219	21.0	58	22	AAE18913	Peptide #4188 enco
28	219	21.0	58	23	ABG37397	Human peptide enco
29	219	21.0	58	23	ABG39657	Human peptide enco
30	174	16.7	56	22	ABB28178	Human peptide #829
31	174	16.7	56	22	ABB33353	Peptide #859 enco
32	174	16.7	56	22	ABB18813	Protein #812 enco
33	174	16.7	56	22	AAE18913	Human brain expres
34	174	16.7	56	22	AAE18913	Human bone marrow
35	174	16.7	56	22	AAE18913	Peptide #839 enco
36	174	16.7	56	22	AAE18913	Peptide #855 enco
37	174	16.7	56	22	AAE18913	Peptide #814 enco
38	174	16.7	56	23	ABG36184	Human peptide enco
39	133	12.8	25	22	AAO8780	Human polypeptide
40	101	9.7	454	17	AAE18913	Human Yes-associat
41	101	9.7	454	22	ABG14507	Novel human diagno
42	90.5	8.7	434	21	AAE18913	Breast and ovarian
43	89	8.5	777	23	ABP41839	Human ovarian anti
44	89	8.5	874	22	ABP41839	Drosophila melanog
45	87	8.4	732	23	ABB93616	Herbicidally activ

ALIGNMENTS

RESULT 1
AAE18913
ID AAE18913: standard; Protein; 197 AA.

AC AAE18913;
DT 17-MAY-2002. (first entry)
XX Mouse MCIP associated proein #4.
DE Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;
KW heart failure; cardiomyopathy; heart disease; mouse.
XX Mus musculus.
XX WO20020491-A2.
XX 17-JAN-2002.
XX 06-JUL-2001; 2001WO-US21662.
XX 07-JUL-2000; 2000US-216601P.
XX 13-FEB-2001; 2001US-0782953.
XX (TEXA) UNIV TEXAS SYSTEM.
XX (WILL/) WILLIAMS S R.
XX (ROTH/) ROTHERMEL B.
XX Williams SR, Rothermel B;
XX WPI, 2002-179698/23.
XX N-P5DB; AAD30154.

PT Screening for modulators of muscle calcineurin interacting protein
 PT (MCIP) binding, expression or phosphorylation, useful for treating
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,
 PT calcineurin and a test compound -

XX Disclosure: Page 154, 174pp; English.

XX The invention relates to muscle calcineurin interacting proteins (MCIPs)
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical
 CC complex with the catalytic subunit of calcineurin and increased levels
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate
 CC transcription of certain target genes. The invention also relates to
 CC methods for identifying modulators of MCIP binding, expression or
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin
 CC may be used for treating cardiac hypertrophy and heart failure.
 CC Antibodies to MCIP can be used in characterising the MCIP content of
 CC healthy and diseased tissues and subsequently for determining the
 CC presence or absence of cardiomyopathy or as predictor of heart disease.
 CC The present sequence is mouse MCIP associated protein.
 CC Note: This sequence has been described as mouse MCIP2 encoding DNA in
 CC the specification, however the sequence seems to be a MCIP associated
 CC protein.

XX Sequence 197 AA;

Query Match 100.0%; Score 1041; DB 23; Length 197;

Best Local Similarity 100.0%; Pred. No. 7e-104;
 Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPAPSDCDVSTLVACVVDVEVTNOEVEKEFGELFRITDECTFOLFSPFRVRINFESH 60

DB 1 MPAPSDCDVSTLVACVVDVEVTNOEVEKEFGELFRITDECTFOLFSPFRVRINFESH 60

QY 61 PKSAARARIEHETOPRGKKLKYFAOVQTPETDGDKLHLAPPQAKQFLISPPSSPPYG 120

DB 61 PKSAARARIEHETOPRGKKLKYFAOVQTPETDGDKLHLAPPQAKQFLISPPSSPPYG 120

QY 121 WKESIDATPVLYNDLLYAAVAKLGEGEKYELHAGTESTPSVVVHVCSDMEEDDPKTSRK 180

DB 121 WKESIDATPVLYNDLLYAAVAKLGEGEKYELHAGTESTPSVVVHVCSDMEEDDPKTSRK 180

QY 181 PKTIQRRRLPPSVSN 197

DB 181 PKTIQRRRLPPSVSN 197

RESULT 2

AAE18916
 ID AAE18916 standard; Protein; 192 AA.

AC AAE18916;

DT 17-MAY-2002 (first entry)

DE Human MCIP associated protein #3.

KW Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;

KW heart failure; cardiomyopathy; heart disease; human.

OS Homo sapiens.

PN WO200204491-A2.

PD 17-JAN-2002.

PF 06-JUL-2001; 2001WO-US21662.

PR 07-JUL-2000; 2000US-216601P.

PR 13-FEB-2001; 2001US-0782953.

PA (TEXA) UNIV TEXAS SYSTEM.

PA (WILL/) WILLIAMS S R.

PA (ROTH/) ROTHERMEL B.

XX Williams SR, Rothermel B;

XX WPI: 2002-179698/23.

DR N-PSDB; AAD30157.

PT Screening for modulators of muscle calcineurin interacting protein
 PT (MCIP) binding, expression or phosphorylation, useful for treating
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,
 PT calcineurin and a test compound -

XX Example 1; Page 165-166; 174pp; English.

XX The invention relates to muscle calcineurin interacting proteins (MCIPs)
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical
 CC complex with the catalytic subunit of calcineurin and increased levels
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate
 CC transcription of certain target genes. The invention also relates to
 CC methods for identifying modulators of MCIP binding, expression or
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin
 CC may be used for treating cardiac hypertrophy and heart failure.
 CC Antibodies to MCIP can be used in characterising the MCIP content of
 CC healthy and diseased tissues and subsequently for determining the
 CC presence or absence of cardiomyopathy or as predictor of heart disease.
 CC The present sequence is human MCIP associated protein.
 CC Note: This sequence has been described as human MCIP splice variant in
 CC the specification, however the sequence seems to be a MCIP associated
 CC protein.

XX Sequence 192 AA;

Query Match 94.1%; Score 980; DB 23; Length 192;

Best Local Similarity 96.4%; Pred. No. 2.5e-97;
 Matches 185; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 6 MDCDSTLVACVVDVEVTNOEVEKEFGELFRITDDCTFOLFSPFRVRINFSPKSA 65

DB 1 MDCDSTLVACVVDVEVTNOEVEKEFGELFRITDDCTFOLFSPFRVRINFSPKSA 60

QY 66 RARIEHETOPRGKKLKYFAOVQTPETDGDKLHLAPPQAKQFLISPPSSPPVGMKPS 125

DB 61 RARIEHETOPRGKKLKYFAOVQTPETDGDKLHLAPPQAKQFLISPPSSPPVGMQPIN 120

QY 126 DATPVLYNDLLYAAVAKLGEGEKYELHAGTESTPSVVVHVCSDMEEDDPKTSRK 185

DB 121 DATPVLYNDLLYAAVAKLGEGEKYELHAGTESTPSVVVHVCSDMEEDDPKTSRK 180

QY 186 TRRRLPPSVSN 197

DB 181 TRRRLPPSVSN 192

RESULT 3

AAB31788
 ID AAB31788 standard; Protein; 234 AA.

AC AAB31788;

DT 30-APR-2001 (first entry)

DE Amino acid sequence of a human detoxification protein.

KW Human; detoxification protein; DCTX; cancer; leukemia; melanoma;

KW adenocarcinoma; autoimmune disorder; inflammatory disorder;

KW rheumatoid arthritis; asthma; atherosclerosis; autoimmune thyroiditis;

KW psoriasis; ulcerative colitis; infection; cell proliferative disorder;

KW actinic keratosis; arteriosclerosis; cirrhosis; hepatitis.

OS Homo sapiens.

PN Key

FT Peptide 1..27
 FT Location/Qualifiers
 FT /note= "signal peptide"

CC presence or absence of cardiomyopathy or as predictor of heart disease.
 CC The present sequence is human MCIP associated protein.
 CC Note: This sequence has been described as murine splice variant in
 CC the specification, however the sequence seems to be a MCIP associated
 CC protein.

XX Sequence 197 AA;

Query Match 61.6%; Score 641; DB 23; Length 197;
 Best Local Similarity 64.2%; Pred. No. 9.2e-61;
 Matches 120; Conservative 29; Mismatches 30; Indels 8; Gaps 2;

QY 11 STLAVCVDEVEFTNOEVEKEFLPTDYECVTFQFKSFRVRINFSHPKSAARARIE 70
 DB 11 SSLIACVANSDFSESETRAKFESLFRDYKDIITFOYFKSFRVRINFSNPFSAADARLQ 70
 QY 71 LHETQFRGKKLYFAQVOTPETDGDKLHAPPQAPAKOFLISPPSPVGVKPISDATPV 130
 DB 71 LHETQFRGKKLYFAQVOTPETDGDKLHAPPQAPAKOFLISPPSPVGVKPISDATPV 130
 QY 131 LNYDLLYAVAKLGPGEKYLHAGTSTPSVHVHVCDSMDEEEDPKTS----PKPKIIQT 186
 DB 127 INYDLLYAVAKLGPGEKYLHAGTSTPSVHVHVCDSMDEEEDPKTS----PKPKIIQT 186
 QY 187 RRCLPP 193
 DB 187 RRPETP 193

RESULT 7

AAE18918
 ID AAE18918 standard; Protein; 212 AA.

XX AAE18918;
 AC AAE18918;

DT 17-MAY-2002 (first entry)

DE Human MCIP associated protein #5.

XX Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;
 KW heart failure; cardiomyopathy; heart disease; human.
 XX Homo sapiens.

XX WO200204491-A2.
 XX 17-JAN-2002.

XX 06-JUL-2001; 2001WO-US21662.
 XX 07-JUL-2000; 2000US-216601P.
 PR 13-FEB-2001; 2001US-0782953.

XX (TEXA) UNIV TEXAS SYSTEM.
 PA (WILL/) WILLIAMS S R.
 PA (ROTH/) ROTHERMEL B.
 XX Williams SR, Rothermel B;

XX WPI; 2002-179698/23.
 DR N-PSDB; AAD30159.

XX Screening for modulators of muscle calcineurin interacting protein
 PT (MCIP) binding, expression or phosphorylation, useful for treating
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,
 PT calcineurin and a test compound -

PS Disclosure; Page 171-172; 174pp; English.

XX The invention relates to muscle calcineurin interacting proteins (MCIPs)
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical
 CC complex with the catalytic subunit of calcineurin and increased levels
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate

CC transcription of certain target genes. The invention also relates to
 CC methods for identifying modulators of MCIP binding, expression or
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin
 CC may be used for treating cardiac hypertrophy and heart failure.
 CC Antibodies to MCIP can be used in characterising the MCIP content of
 CC healthy and diseased tissues and subsequently for determining the
 CC presence or absence of cardiomyopathy or as predictor of heart disease.
 CC The present sequence is human MCIP associated protein described in the
 CC invention.

XX Sequence 212 AA;

Query Match 60.9%; Score 633.5; DB 23; Length 212;
 Best Local Similarity 67.9%; Pred. No. 6.5e-60;
 Matches 127; Conservative 17; Mismatches 32; Indels 11; Gaps 2;

QY 11 STLAVCVDEVEFTNOEVEKEFLPTDYECVTFQFKSFRVRINFSHPKSAARARIE 70
 DB 28 TSLFACSVHEAVFEAREQERFEALFTYDDQVTFQFKSFRVRINFSKPEAARARIE 87

QY 71 LHETQFRGKKLYFAQVOTPETDGDKLHAPPQAPAKOFLISPPSPVGVKPISDATPV 130
 DB 88 LHETDNGQKLYFAQVOTPETDGDKLHAPPQAPAKOFLISPPSPVGVKQSEDAMPV 137

QY 131 LNYDLLYAVAKLGPGEKYLHAGTSTPSVHVHVCDSMDEEEDPKTSKPCKIIQTRPG 190
 DB 138 INYDLLCAVSKLGPGEKYLHAGTSTPSVHVHVCSETEEEETK-NPKQKIAQTRPD 196

QY 191 LPPSVSN 197

DB 197 PPTAALN 203

RESULT 8

AAE18912
 ID AAE18912 standard; Protein; 198 AA.

XX AAE18912;
 AC AAE18912;

DT 17-MAY-2002 (first entry)

DE Mouse MCIP associated protein #3.

XX Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;
 KW heart failure; cardiomyopathy; heart disease; mouse.

XX Mus musculus.

XX WO200204491-A2.

XX 17-JAN-2002.

XX 06-JUL-2001; 2001WO-US21662.

XX 07-JUL-2000; 2000US-216601P.

PR 13-FEB-2001; 2001US-0782953.

XX (TEXA) UNIV TEXAS SYSTEM.

PA (WILL/) WILLIAMS S R.

PA (ROTH/) ROTHERMEL B.

XX Williams SR, Rothermel B;

XX WPI; 2002-179698/23.

DR N-PSDB; AAD30153.

XX Screening for modulators of muscle calcineurin interacting protein
 PT (MCIP) binding, expression or phosphorylation, useful for treating
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,
 PT calcineurin and a test compound -

PS Disclosure; Page 151-152; 174pp; English.

CC The invention relates to muscle calcineurin interacting proteins (MCIPs)
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical
 CC complex with the catalytic subunit of calcineurin and increased levels
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate
 CC transcription of certain target genes. The invention also relates to
 CC methods for identifying modulators of MCIP binding, expression or
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin
 CC may be used for treating cardiac hypertrophy and heart failure.
 CC Antibodies to MCIP can be used in characterizing the MCIP content of
 CC healthy and diseased tissues and subsequently for determining the
 CC presence or absence of cardiomyopathy or as predictor of heart disease.
 CC Note: This sequence has been described as human MCIP in the
 CC specification, however the sequence seems to be a MCIP associated
 CC protein.
 CC
 SQ Sequence 198 AA;
 Query Match 60.8%; Score 632.5; DB 23; Length 198;
 Best Local Similarity 64.4%; Pred. No. 7.6e-60;
 Matches 121; Conservative 24; Mismatches 34; Indels 9; Gaps 2;
 QY 11 STLVACVVDVEVFTNOEVEKEFEGLFRTYDECVTPLFKSPFRVRINFSHPKSAARAE 70
 DB 11 SSLICVANDVPSSESTRAKESLFRTYDKDTTFQYFKSPKRVINFSNPUSAADARLR 70
 QY 71 LHETQFRGKKLKYFAOVOTPETDGDKLHLPAPQAKOFLISPPSSPPVGMKPISDATPV 130
 DB 71 LKTEFLGKEMKLYFAQTLHGSS---HLAPNPDKOFLISPPSPVGMQVEDATPV 126
 QY 131 LNYDLLVAVALKGPEKELYLAGTESTPSVVVHVCDSDMEEDPK-----TSPKIIQ 185
 DB 127 INYDLLVAISKLGPEKELYLAATDTPSPVVHVCDSDNEEBEEMERMRPKPKIIQ 186
 QY 186 TRRPGLP 193
 DB 187 TRPEYTP 194
 RESULT 9
 AAE18915
 ID AAE18915 standard; Protein; 197 AA.
 XX
 AC AAE18915;
 XX
 DT 17-MAY-2002 (first entry)
 XX
 DE Human MCIP associated protein #2.
 XX
 KM Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;
 XX heart failure; cardiomyopathy; heart disease; human.
 XX
 OS Homo sapiens.
 XX
 PN WO200204491-A2.
 XX
 PD 17-JAN-2002.
 XX
 PF 06-JUL-2001; 2001WO-US21662.
 XX
 XX 07-JUL-2000; 2000US-216601P.
 PR 13-FEB-2001; 2001US-0782953.
 XX
 PA (TEXA) UNIV TEXAS SYSTEM.
 PA (WILL/) WILLIAMS S R.
 PA (ROTH/) ROTHERMEL B.
 XX
 PI Williams SR, Rothermel B;
 XX
 DR MPI: 2002-179698/23.
 DR N-PSDB; AAD30156.
 XX
 PT Screening for modulators of muscle calcineurin interacting protein

PT (MCIP) binding, expression or phosphorylation, useful for treating
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,
 PT calcineurin and a test compound -
 XX
 XX Example 1; Page 161-162; 174pp; English.
 CC The invention relates to muscle calcineurin interacting proteins (MCIPs)
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical
 CC complex with the catalytic subunit of calcineurin and increased levels
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate
 CC transcription of certain target genes. The invention also relates to
 CC methods for identifying modulators of MCIP binding, expression or
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin
 CC may be used for treating cardiac hypertrophy and heart failure.
 CC Antibodies to MCIP can be used in characterizing the MCIP content of
 CC healthy and diseased tissues and subsequently for determining the
 CC presence or absence of cardiomyopathy or as predictor of heart disease.
 CC Note: This sequence has been described as murine MCIP splice variant in
 CC the specification, however the sequence seems to be a MCIP associated
 CC protein.
 CC
 SQ Sequence 197 AA;
 Query Match 59.8%; Score 622; DB 23; Length 197;
 Best Local Similarity 63.6%; Pred. No. 1e-58;
 Matches 119; Conservative 26; Mismatches 34; Indels 8; Gaps 2;
 QY 11 STLVACVVDVEVFTNOEVEKEFEGLFRTYDECVTPLFKSPFRVRINFSHPKSAARAE 70
 DB 11 SATIACHDPRVFDGLORAFESLFRTYDKDTTFQYFKSPKRVINFSNPUSAADARLQ 70
 QY 71 LHETQFRGKKLKYFAOVOTPETDGDKLHLPAPQAKOFLISPPSSPPVGMKPISDATPV 130
 DB 71 LKTEFLGKEMKLYFAQTLHGSS---HLAPNPDKOFLISPPSPVGMQVEDATPV 126
 QY 131 LNYDLLVAVALKGPEKELYLAGTESTPSVVVHVCDSDMEEDPKS---PKPKIIQ 186
 DB 127 INYDLLVAISKLGPEKELYLAATDTPSPVVHVCDSDNEEBEEMERMRPKPKIIQ 186
 QY 187 RRPGLP 193
 DB 187 RRPETYP 193
 RESULT 10
 AAE18911
 ID AAE18911 standard; Protein; 198 AA.
 XX
 AC AAE18911;
 XX
 DT 17-MAY-2002 (first entry)
 XX
 DE Human MCIP associated protein #2.
 XX
 KM Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;
 XX heart failure; cardiomyopathy; heart disease; human.
 XX
 OS Homo sapiens.
 XX
 PN WO200204491-A2.
 XX
 PD 17-JAN-2002.
 XX
 PF 06-JUL-2001; 2001WO-US21662.
 XX
 XX 07-JUL-2000; 2000US-216601P.
 PR 13-FEB-2001; 2001US-0782953.
 XX
 PA (TEXA) UNIV TEXAS SYSTEM.
 PA (WILL/) WILLIAMS S R.
 PA (ROTH/) ROTHERMEL B.
 XX

XX Disclousure; SEQ ID NO 41193; 21pp + Sequence Listing; English.
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB16175), expressed DNA
 CC sequences (AB16173-AB16172).
 CC (AB57737-AB57202).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WFO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX Sequence 292 AA;
 SO
 Query Match 38.8%; Score 404; DB 22; Length 292;
 Best Local Similarity 43.2%; Pred. No. 5.6e-35;
 Matches 83; Conservative 35; Mismatches 62; Indels 12; Gaps 4;
 QY 4 PSMDD- - - - - VSTLVACVDEVEFTNOEYKKEFEGLFRTYDECVTFOLEKSFRRVRINF 58
 DB 104 PEVDADSFDDLPSTIIVINIHSEVFANPELKIAMELELFTFSBSATFQWLRGFRRLRVNY 163
 QY 59 SHPKSARARIELHETQFRGKK-LKLYFAOVQTPETDGDKLHLAPQAPQFLISPPSP 117
 DB 164 DNAIAAANRILKHQYEFKKTIVITCYFAQPTPVSN--XNLQPPAPKQFLISPPSP 220
 QY 118 PVGWRKPSDATTPLNYDLIYAVAKLGPGEKYLEHAGTSTPSVYVAVCDMEBEDPPT 177
 DB 221 PAGMEPRERGEPLVNHDLALALSLTPGESHLEHPSQSDQPAIIVHTA--MLAETGPG 277
 QY 178 SPKPKIOTRRP 189
 DB 278 QVKAPIVQTKCP 289
 RESULT 13
 AAG01768
 ID AAG01768 standard; Protein; 111 AA.
 XX
 AC AAG01768;
 XX
 DT 06-OCT-2000 (first entry)
 XX
 DE Human secreted protein, SEQ ID NO: 5849.
 XX
 KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping.
 XX
 OS Homo sapiens.
 XX
 PN EP1033401-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 21-FEB-2000; 2000EP-0200610.
 XX
 PR 26-FEB-1999; 99US-0122487.
 XX
 PA (GEST) GENSET.
 XX
 PI Dumas Milne Edwards J, Duclert A, Giordano J;
 XX
 DR WPI; 2000-500361/45.
 DR N-PSDB; AAC01774.
 XX
 PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNA that correspond to 5' ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 XX
 PS Claim 13; SEQ ID 5849; 71pp + CD-ROM; English.

XX The present sequence is a polypeptide encoded by one of a large number
 CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
 CC were prepared from total human RNAs or polyA+ RNAs derived from 30
 CC different tissues. EST sequences usually correspond mainly to the 3'
 CC untranslated region (UTR) of the mRNA because they are often obtained
 CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
 CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
 CC those cases where longer cDNA sequences have been obtained, the full 5'
 CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
 CC ends and can therefore be used to obtain full length cDNAs and genomic
 CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
 CC chromosome mapping procedures. They are used to obtain upstream
 CC regulatory sequences and to design expression and secretion vectors.
 CC
 XX Sequence 111 AA;
 SO
 Query Match 31.0%; Score 323; DB 21; Length 111;
 Best Local Similarity 59.0%; Pred. No. 7.8e-27;
 Matches 62; Conservative 19; Mismatches 20; Indels 4; Gaps 1;
 QY 11 STLVACVDEVEFTNOEYKKEFEGLFRTYDECVTFOLEKSFRRVRINFSPKSAARARIE 70
 DB 11 SSLIACVANSIDIFSESSETRAFESLFTYDKDITFOYFKSFRKRINPSNPSAADARLQ 70
 QY 71 LHETQFRGKKLKLTPAQVQTPETDGDKLHLAPQAPQFLISPPS 115
 DB 71 LHKTEFLGKEMKLYFAQTLHIGSS--HLAPPNPDQFLISPPA 111
 RESULT 14
 AAB58452
 ID AAB58452 standard; Protein; 142 AA.
 XX
 AC AAB58452;
 XX
 DT 14-MAR-2001 (first entry)
 XX
 DE Lung cancer associated polypeptide sequence SEQ ID 790.
 XX
 KW Human; lung cancer associated protein; neuroprotective; cyostatic;
 KW cardioactive; immunomodulatory; muscular active; vulnery;
 KW gastrointestinal; nephrotoxic; antiinfective; gynecological;
 KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
 KW proliferative disorder; wound healing; infectious disease.
 XX
 OS Homo sapiens.
 XX
 PN WO20005180-A2.
 XX
 PD 21-SEP-2000.
 XX
 PF 08-MAR-2000; 2000WO-US05918.
 XX
 PR 12-MAR-1999; 99US-0124270.
 XX
 PA (HUMA-) HUMAN GENOME SCT INC.
 PA (ROSE/) ROSEN C A.
 XX
 PI Ruben SM;
 XX
 DR WPI; 2000-587514/55.
 DR N-PSDB; AAF18328.
 XX
 PT Lung cancer associated gene sequences, referred to as lung cancer
 PT antigens, useful for treatment, prevention, and diagnosis of disorders
 PT such as lung cancer -
 XX
 PS Claim 11; Page 1325-1326; 1425pp; English.
 CC Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer
 CC associated proteins represented in AAB58106 - AAB58548. Lung cancer
 CC associated proteins and polynucleotide sequences, their agonists, and

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 14, 2002, 18:14:56 ; Search time 1655.79 Seconds
(without alignments)
3480.130 Million cell updates/sec

Title: US-09-782-953-7
Perfect score: 1043
Sequence: 1 MDRFDFSYNPSLIICVAND.....RPKPKIIQTRRPEYPIHLS 198

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DBV=xlh
-Q=/cgn2_1/USPTO.spool/US09782953/runat_11122002_114429_17340/app.query.fasta_1.1173
-DB=GenEmbl -OFMT=fastap -SUPPLX=rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bite -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptco -NORW=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09782953 @CNG 1 1 2844 @runat_11122002_114429_17340 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MWAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb.ba:*
2: gb.htg:*
3: gb.in:*
4: gb.om:*
5: gb.ov:*
6: gb.pat:*
7: gb.ph:*
8: gb.pl:*
9: gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sy:*
13: gb.un:*
14: gb.vi:*
15: em.ba:*
16: em.fun:*
17: em.hum:*
18: em.in:*
19: em.mu:*
20: em.om:*
21: em.or:*
22: em.ov:*
23: em.pat:*
24: em.ph:*
25: em.pl:*
26: em.ro:*
27: em.sts:*
28: em.un:*

29: em.vi:*
30: em.htg_hum:*
31: em.htg_inv:*
32: em.htg_other:*
33: em.htg_mus:*
34: em.htg_pln:*
35: em.htg_rtd:*
36: em.htg_mam:*
37: em.htg_vrt:*
38: em.sy:*
39: em.htgo_hum:*
40: em.htgo_mus:*
41: em.htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1043	100.0	597	6	AX365315	Sequence
2	1043	100.0	597	10	AF237790	Mus muscu
3	1028	98.6	2198	10	BC013551	Mus muscu
4	1025	98.3	2141	10	AF263239	Mus muscu
5	995.5	95.4	626	10	AB075973	Rattus no
6	992.5	95.2	2216	10	CGU60263	Cricetus
7	970.5	93.0	2331	6	AX365321	Sequence
8	970.5	93.0	2346	9	HSU85267	Homo sapien
9	957.5	91.8	2348	6	AX281651	Sequence
10	920.5	88.3	597	6	AX365312	Sequence
11	920.5	88.3	597	10	AF237789	Mus muscu
12	912.5	87.5	621	10	AF282255	Mus muscu
13	912.5	87.5	2125	10	AF260717	Mus muscu
14	890	85.3	2224	10	AF263240	Mus muscu
15	875	83.9	2212	6	AX365324	Sequence
16	875	83.9	2227	9	HSU85366	Homo sapien
17	875	83.9	2289	9	BC002864	Homo sapi
18	875	83.9	2407	9	AK092184	Homo sapi
19	868	83.2	599	6	AX365311	Sequence
20	850.5	81.5	2284	9	HSU85265	Homo sapien
21	847.5	81.3	2173	6	AX410694	Sequence
22	847.5	81.3	2173	9	HSU28833	Homo sapien
23	844.5	81.0	2174	6	AR034241	Sequence
24	844.5	81.0	2579	9	AK055845	Homo sapi
25	839.5	80.5	798	9	AF400429	Homo sapi
26	807	77.4	562	9	HSU53821	Homo sapien
27	634.5	60.8	3295	10	AB061525	Mus muscu
28	632.5	60.6	594	6	AX365318	Sequence
29	632.5	60.6	594	10	AF237791	Mus muscu
30	632.5	60.6	599	10	AF237887	Mus muscu
31	627.5	60.2	3227	10	AB061524	Mus muscu
32	623	59.7	934	6	AX074351	Sequence
33	623	59.7	3159	6	AX420425	Sequence
34	623	59.7	3261	9	AK090990	Homo sapi
35	615.5	59.0	3184	6	AX329596	Sequence
36	615.5	59.0	3184	6	AX333015	Sequence
37	615.5	59.0	3184	6	AX365327	Sequence
38	615.5	59.0	3184	6	AX420436	Sequence
39	615.5	59.0	3184	9	HUMZAKI4	mRNA
40	610	58.5	3240	9	AY034085	Homo sapi
41	610	58.5	3253	9	AY034086	Homo sapi
42	594.5	57.0	776	10	AF237888	Mus muscu
43	592.5	56.8	828	6	AX365330	Sequence
44	592.5	56.8	828	6	AX420437	Sequence
45	592.5	56.8	828	9	AF176116	Homo sapi

ALIGNMENTS

RESULT 1

AX365315
 LOCUS AX365315 597 bp DNA linear PAT 15-FEB-2002
 DEFINITION Sequence 5 from Patent WO0204491.
 ACCESSION AX365315
 VERSION AX365315.1 GI:18697045
 KEYWORDS
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 1 Williams, S.R. and Rothermel, B.
 Methods and compositions relating to muscle selective calcineurin interacting protein (mcip)
 Patent: WO 0204491-A 5 17-JAN-2002;
 Board of Regents, The University of Texas System (US); Williams, Sanders R. (US); Rothermel, Beverly (US)

FEATURES
 source
 1..597
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 1..597
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD23809.1"
 /db_xref="GI:18697046"
 /translation="MDPRDSYNFSSLIACVANDVSESETRAKFSLPTDYDKDTT
 FOYKSKRVRINFSNPLSADARLHKTEFLGKMKLYAOTLHIGSSHLAPNPD
 KQPLISPPASPVGMKQVEDATPVINDLYALSKLGGKYEELHAATDPTPSVVAV
 CESDQENEEEEEEMERKRPKIIQTRRPEYPIHLS"

CDS
 BASE COUNT 170 a 156 c 142 g 129 t
 ORIGIN

Alignment Scores:
 Pred. No.: 2,92e-96 Length: 597
 Score: 1043.00 Matches: 198
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 0

US-09-782-953-7 (1-198) x AX365315 (1-597)

QY 1 MetAappheaRgaspPheSerTyRanPheSerSerLeuIleAlaCyValAlaAsnAp 20
 Db 1 ATGCAATTTTAGAGGACTTTAGCTACAAATTTAGCTCCCTGATGCTGTCGCAACGAT 60
 QY 21 AppValPheSerGluSerGluThrArgAlaAlaPheGluSerLeuPheArgThTyRasp 40
 Db 61 GATGCTTCAGCGAAAGTAGACACAGGCGCAAAATTTGAATCCCTCTTCAGAACATATAC 120
 QY 41 LybAapThThPhagIntyRPhelVserPheLyserValArgIleAsnPheseraen 60
 Db 121 AAGGACACCACTTCAGATTTTAAAGCTTCACACGTCGCGATTAACCTTCAGCAAC 180
 QY 61 ProLeuSerAlaAlaAspAlaArgLeuAArgLeuAlaVleThGluPheLeuGlu 80
 Db 181 CCTTATTCGACGCCATCCAGGCTCGGCTGCACAAACGAGTCTCTGGGGAAGGAA 240
 QY 81 MetLyLeuTyRPhelAglInThLeuHsIleGlySerSerHsIleuAlaProProan 100
 Db 241 ATGAAGTTGATTTGCTCAGACTTACACATAGAAATTCACACCTGCTCGCCCAAT 300
 QY 101 ProAapLyGlnPheLeuIleSerProProAlaSerProProValGlyTrpLyGlnVal 120
 Db 301 CCCGACAAACAGTCTCTCATCTCCCTCCGCTCTCTCCGCTGCGTGGAAACAGTA 360
 QY 121 GluAapAlaThProValIleAenTyRaspLeuTyAlaIleSerTyLeuGlyPro 140
 Db 361 GAAGATGCCACCCCTCATTAATTCAGATCTTTATATGCCATCTCCAAAGCTGGGCCA 420
 QY 141 GtGtLbLyTyGtLbLeuHsAlaAlaThRaspProThProSerValValHsVal 160

Db 421 GGAGAGAAAGTATGACTGCATGCAGCAGACACCCCACTCCAGTGTGTGTCACAGTG 480
 QY 161 CysGluSerAapGlnGluAangluGluGluGluGluMetGluArgMetLyserPro 180
 Db 481 TGTGAGAGTGCACAAAGAGATGAGAGAGAAAGAGAGATGAGAGATGAAGAGACCC 540
 QY 181 LybProLyValIleIleGlnThRArgArgProGluTyRThProIleHsIleAsnSer 198
 Db 541 AAGCCCAAAATCATCTCAGACACGAGACCGAGTACACACCATTCACCTTAGC 594

RESULT 2
 AF237790
 LOCUS AF237790 597 bp mRNA linear ROD 12-APR-2000
 DEFINITION Mus musculus myocyte-enriched calcineurin interactin protein 1
 ACCESSION AF237790
 VERSION AF237790.1 GI:7542528
 KEYWORDS
 SOURCE Mus musculus.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 1 (bases 1 to 597)
 Rothermel, B.A., Vega, R.B., Yang, J., Wu, H., Bassel-Duby, R.S. and Williams, R.S.
 A protein encoded within the Down Syndrome Critical Region is enriched in striated muscles and inhibits calcineurin signaling
 J. Biol. Chem. (2000) In press
 2 (bases 1 to 597)
 Rothermel, B.A., Vega, R.B., Yang, J., Wu, H., Bassel-Duby, R.S. and Williams, R.S.
 Direct Submission
 Submitted (23-FEB-2000) Internal Medicine, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, TX 75390, USA

FEATURES
 source
 1..597
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /chromosome="16"
 1..597
 /note="similar to the EST with GenBank Accession Number AA200984, similar to Homo sapiens DSCR1"

CDS
 BASE COUNT 170 a 156 c 142 g 129 t
 ORIGIN

Alignment Scores:
 Pred. No.: 2,92e-96 Length: 597
 Score: 1043.00 Matches: 198
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 0

US-09-782-953-7 (1-198) x AF237790 (1-597)

QY 1 MetAappheaRgaspPheSerTyRanPheSerSerLeuIleAlaCyValAlaAsnAp 20
 Db 1 ATGCAATTTTAGAGGACTTTAGCTACAAATTTAGCTCCCTGATGCTGTCGCAACGAT 60
 QY 21 AppValPheSerGluSerGluThrArgAlaAlaPheGluSerLeuPheArgThTyRasp 40
 Db 61 GATGCTTCAGCGAAAGTAGACACAGGCGCAAAATTTGAATCCCTCTTCAGAACATATAC 120

QY 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60
 Db 121 AAGGACACACCTTCAGTATTTAAGAGCTTCAACAGTGTCCGGATAACTTCAGCAAC 180
 QY 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
 Db 181 CCCATTATCTGACGCCGATCCAGGCTGGGCTGCCAAGAGCCGAGTTCCTGGGAAGGAA 240
 QY 81 MetLysLeuTyrPheAlaGlnThrLeuHisLysGlySerSerHisLeuAlaProProAsn 100
 Db 241 ATGAAGTGTATTTGCTCAGACTTTACACATAGGAAGTTCACACCTGGCTCCGCCCAAT 300
 QY 101 ProAspLysGlnPheLeuLysSerProProAlaSerProProValGlyTyrLysGlnVal 120
 Db 301 CCCGACAAACAGTTCCTCATCTCCCTCCGGCTCTCTCCCTGGCTGGCTGGCAACAGTA 360
 QY 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro 140
 Db 361 GAAGATGCCACCCCGCTCATTAATACGATCTTTATATGCACTTCCAAAGCTGGGGCCA 420
 QY 141 GlyLysLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160
 Db 421 GGAGAGAAGTATGAATCATGTCATGCGAGCAGACACCCACTCCCACTGTGTGTCCACGTG 480
 QY 161 CysGluSerAspGlnGluAsnGluGluGluGluMetGluArgMetLysArgPro 180
 Db 481 TGTGAGAGTGACCAAGAGAATGAGGAGGAAGAGGAGATGGAGAGATGAAGAGACCC 540
 QY 181 LysProLysIleGlnThrArgArgProGluTyrThrProIleHisLeuSer 198
 Db 541 AAGCCCAAAATCATCCAGACAGGAGCCGAGTACACACCGATCCACCTTAGC 594

RESULT 3
 BC013551 2198 bp mRNA linear ROD 07-AUG-2002
 LOCUS Mus musculus, Down syndrome critical region homolog 1 (human),
 DEFINITION Clone MGC:19348 IMAGE:4236038, mRNA, complete cds.
 ACCESSION BC013551
 VERSION BC013551.1 GI:15488840
 KEYWORDS MGC.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 2198)
 Strausberg,R.
 Direct Submission
 Submitted (04-SEP-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapsb-r@mail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>
 Contact: (Dickson, Mark) md@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 25 Row: m Column: 6.

FEATURES

Location/Qualifiers
 1..2198
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /map="FVB/N"
 /clone="MGC:19348 IMAGE:4236038"

/tissue_type="kidney, normal. 5 month old male mouse."
 /clone_lib="NCI_CGAP_Kid14"
 /lab_host="DH10B"
 /notes="Vector: pCMV-SPORT6"
 77..673
 /codon_start=1
 /product="Down syndrome critical region homolog 1 (human)"
 /protein_id="AAH13551.1"
 /db_xref="GI:15488841"
 /db_xref="LocusID:54720"
 /translation="MHFRDYSYFSSLIACVANDDVSESETRAKFESLFRVYDKDTT
 FOYFKSFRVIRINFSNPLISAADARLRLHKTFLKEMKLYFACTLHIGSSHLAPNPD
 KQFLISPPASPVGNKQVEDATPVINDLYALSKLGPGEKYLHNATDTTTSVVVHV
 CESDQENBEBEEMERMRPKPKIQTRRPEYPIHLS"

BASE COUNT 533 a 539 c 569 g 557 t
 ORIGIN

Alignment Scores:

Pred. No.: 4-67e-94 Length: 2198
 Score: 1028.00 Matches: 196
 Percent Similarity: 98.93% Conservative: 0
 Best Local Similarity: 98.99% Mismatches: 2
 Query Match: 98.56% Indels: 0
 DB: 10 Gaps: 0

US-09-782-953-7 (1-198) x BC013551 (1-2198)

QY 1 MetAspPheArgAspPheSerTyrAsnPheSerSerLeuIleAlaCysValAlaAsnAsp 20
 Db 77 ATGCATTTTAGGGACTTTAGCTACAATTTTAGCTCCCTGATTGTTGTGGCAACGAT 136
 QY 21 AspValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40
 Db 137 GATGCTTTCAGCAAGTAGGACCGAGGCGCAATTTGAATCCCTCTTCAGAACATATGAC 196
 QY 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60
 Db 197 AAGGACACCACTTCCAGTATTTAAGAGCTTCAACAGTGTCCGGATAACTTCAGCAAC 256
 QY 61 ProLeuSerAlaAlaAspAlaArgLeuHisLysThrGluPheLeuGlyLysGlu 80
 Db 257 CCTTATCTGCGAGCCGATGCCAGGCTGCGCTGCACAGACCGAGTTCCTGGGAAGGAA 316
 QY 81 MetLysLeuTyrPheAlaGlnThrLeuHisLysGlySerSerHisLeuAlaProProAsn 100
 Db 317 ATGAAGTGTATTTGCTCAGACTTTTACATAGGAAGTTCACACTGGCTCCGCCCAAT 376
 QY 101 ProAspLysGlnPheLeuLysSerProProAlaSerProProValGlyTyrLysGlnVal 120
 Db 377 CCGGACAAACAGTTCCTCATCTCCCTCCGGCTCTCTCCCTGGCTGGGAACAGTA 436
 QY 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro 140
 Db 437 GAAGATGCCACCCCGCTCATAAATTCAGATCTTTTATATGCCATCTCCAAGCTGGGGCA 496
 QY 141 GlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160
 Db 497 GGAGAGAAGTATGAATCATGTCATGCGAGCAGACACCACTTCACAGTGTGTGTCCACGTG 556
 QY 161 CysGluSerAspGlnGluAsnGluGluGluMetGluArgMetLysArgPro 180
 Db 557 TGTGAGAGTGACCAAGAGAATGAGGAGGAAGAGATGGAGAGATGAAGAGACCC 616
 QY 181 LysProLysIleGlnThrArgArgProGluTyrThrProIleHisLeuSer 198
 Db 617 AAGCCCAAAATCATCCAGACAGGAGCCGAGTACACACCGATCCACCTCAGC 670

SOURCE

RESULT 4
 AF263239 2141 bp mRNA linear ROD 30-MAY-2000
 LOCUS Mus musculus calcineurin inhibitor mRNA, complete cds,
 DEFINITION alternatively spliced.
 ACCESSION AF263239

VERSION	AF263239.1	GI:8102011
KEYWORDS	Mus musculus.	
SOURCE	Mus musculus.	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Scurionathii; Muridae; Murinae; Mus.	
REFERENCE	1 (bases 1 to 2141)	
AUTHORS	Fuentes,J.J., Pritchard,M.A., Pucharcos,C. and Estivill,X.	
TITLE	Down syndrome candidate region 1 (Dscr1), one of three alternatively spliced exon 1 transcripts	
JOURNAL	Unpublished	
REFERENCE	2 (bases 1 to 2141)	
AUTHORS	Fuentes,J.J., Pritchard,M.A., Pucharcos,C. and Estivill,X.	
TITLE	Direct Submission	
JOURNAL	Submitted (02-MAY-2000) Centre de Genetica Medica i Molecular, IFO, Avia, Castelldefels Km. 2,7, L'Hospital de Llobregat, Barcelona 08907, Spain	
FEATURES	Location/Qualifiers	
SOURCE	1..2141	
	/organism="Mus musculus"	
	/db_xref="taxon:10090"	
	/chromosome="16"	
CDS	31..627	
	/note="Down syndrome candidate region 1; Dscr1; alternatively spliced exon 1"	
	/codon_start=1	
	/product="calcineurin inhibitor"	
	/protein_id="AA072701.1"	
	/db_xref="GI:8102012"	
	/translation="MHFRDYSNFSLLACVANDVDFSESETRAKFESLFRYYDKDTTFOYFSPKRVIRINFSNPLSADARLRKTEBLGKEMLYFAQTILHIGSSHLAPNPPIKQPLSPASPVGMKQVEDAPVINYDLVLAISLGGEXELHAADTTPSVVHHCESDENSEBEREMRMRPKKIIQTRAPERVPHLS"	
BASE COUNT	505 a	527 c 558 g 548 t 3 others
ORIGIN		
Alignment Scores:		
Pred. No.:	9 08e-94	length: 2141
Score:	1025.00	Matches: 195
Percent Similarity:	98.99%	Conservative: 1
Best Local Similarity:	98.48%	Mismatches: 2
Query Match:	98.27%	Indels: 0
DB:	10	Gaps: 0
US-09-782-953-7 (1-198) x AF263239 (1-2141)		
Qy	1 MetAspPheArGspPheSerTYrAsnPhSeSeSerLeuIleAlaCyValAlaAsnAsp	20
Db	31 ATGCATTATTAGGACATAGCTACATCAATTAGCTCCCTGATGCTGTGGCAACGAT	90
Qy	21 AspValPheSeGluSeGluThrArGAlaLysPheGluSerLeuPheArGThrTYrAsp	40
Db	91 GATGCTTCGACCAAGAGTAGACACGAGGCCCAAAATTGAACTCCCTTCAGAACATATGAC	150
Qy	41 LysAspThrArPheGlnTYrPheLysSerPheLysArGValArgIleAsnPheSerAsn	60
Db	151 AAGGACACCCCTTCAGATATTAAAGCTTCAACGTCGCCGATAACTTCAGCAAC	210
Qy	61 ProLeuSerAlaAlaAspAlaArGLeuArGLeuIleLysThrGluPheLysGlyLysGlu	80
Db	211 CCTTATCTGCACGCCGATGCAAGCTCGGCTGACAAAGCCGAGTTCCTGGGAGAGAA	270
Qy	81 MetLysLeuTYrPheAlaGlnThrLysIleIleGlySerSerHisLeuAlaProProAsn	100
Db	271 ATGAAATTGATTGCTCAGACTTTACATAGGAAGTTCAACCTCGGCTCCGCCAAT	330
Qy	101 ProAspLysGlnPheLeuIleSerProProAlaSerProProValGlyTYrLysGlnVal	120
Db	331 CCGACCAAAAGATTCCTCATCTCCCTCCGCTCTCTCCCGTTGGCTGGAAACAAGTA	390
Qy	121 GluAspAlaThrProValIleAsnTYrAspLeuLeuTYrAlaIleSerLysLeuGlyPro	140
Db	391 GAAAGATCCACCCCGCTCATTAATTAGATCTTTTATATGCACTCCCAAGCTGGGCCA	450

[illegible]

```

QY 21 AspValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40
DB 79 GATGTCCTCAGCGAAGTGAGACCAAGGCCAAATTTGAATCCCTCTTCAGGACGTATGAC 138

QY 41 LysAspThrThrPheGlnTyrPheLysSerPheLysAlaValAtgIleAsnPheSerAsn 60
DB 139 AAGGACATCACCCTCCAGTATTTTAAGAGCTTCAAGCGTGTCCGATNAATCTTCAGCAAC 198

QY 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
DB 199 CCCCTGTCTCGCGCGAGCGCCAGGCTGCGGTGCATAAGACAGAGTTCCTGGGAAGGAG 258

QY 81 MetLysLeuTyrPheAlaGluThrLeuHisLysGlySerSerHisLeuAlaProProAsn 100
DB 259 ATGAACCTGTACTTTGCCAGACITTTACATAGGAAGTTTCAACCTCGCTCCGCCCAAT 318

QY 101 ProAspLysGlnPheLeuIleSerProProAlaSerProValGlyTyrLysGlnVal 120
DB 319 CCAGACAAACAGTTTCTCATCTCTCCCTCGCTCTCCACCTGTGTGGCTGGAACAGTA 378

QY 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro 140
DB 379 GAAGACGCTACCCCACTACAACTACAGTCTTTATATGCTATCTCCAAGCTGGGACCA 438

QY 141 GlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160
DB 439 GGAGAGAGTACGAGCTACGCGAGCGACAGACACCAGCCCGCGGTGGTCCACGTG 498

QY 161 CysGluSerAspGln---GluAsnGluGluGluGluMetGluAtgMetLysArg 179
DB 499 TGTGAGAGTACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 558

QY 180 ProlysProLysIleIleGlnThrArgArgProGluTyrThrProIleHisLeuSer 198
DB 559 CCCAAGCCAAAATCATCCAGACTCGGAGGCGAGTACACACCCATCCACCTCAGC 615

RESULT 6
CGU60263 2216 bp mRNA linear ROD 03-SEP-1997
LOCUS Cricetulus griseus Adapt78 (adapt78) mRNA, complete cds.
DEFINITION U60263
VERSION U60263.1 GI:2351390
KEYWORDS Cricetulus griseus.
SOURCE Cricetulus griseus.
ORGANISM Cricetulus griseus.
REFERENCE 1 (bases 1 to 2216)
AUTHORS Crawford,D.R., Leahy,K.P., Abramova,N., Lan,L., Wang,Y. and Davies,K.J.
TITLE Hamster adapt78 mRNA is a Down syndrome critical region homologue that is inducible by oxidative stress
JOURNAL Arch. Biochem. Biophys. 342 (1), 6-12 (1997)
MEDLINE 97329095
PUBMED 9185608
REFERENCE 2 (bases 1 to 2216)
AUTHORS Crawford,D.R., Leahy,K.L. and Davies,K.J.A.
TITLE Direct Submission
JOURNAL Submitted (10-JUN-1996) Biochem. and Mol. Biol., Albany Medical College, 47 New Scotland Avenue, Albany, NY 12208, USA
FEATURES
source 1..2216
/organism="Cricetulus griseus"
/db_xref="taxon:10029"
/cell_type="HA-1"
1..2216
/gene="adapt78"
70..563
/gene="adapt78"
/notes="similar to human DSCR1; an oxidant- and calcium-inducible mRNA"

```

```

/codon_start=1
/product="Adapt78"
/protein_id="AAB68517.1"
/db_xref="GI:2351391"
/translacion="MHFRDNVNFSLIACVANGDVSESETRAKFESLFRTYDKDIT
PQYKSPKRVIRNFNPUSADARLQHLKTEFLGKMKLYFAQTLLHIGSHLAPNPDP
KQFLISPASPVPVQKVEDATPVINYLILYAIKLGPGKEVLELHAATDTTPSVVHV
CESDQNEEEEMERMKRPKIIITRRPEYTPHLS"
BASE COUNT 544 a 515 c 564 g 593 t
ORIGIN
Alignment Scores:
Pred. No.: 1.79e-90 Length: 2216
Score: 992.50 Matches: 191
Percent Similarity: 97.47% Conservative: 2
Best Local Similarity: 96.46% Mismatches: 4
Query Match: 95.16% Indels: 1
DB: 10 Gaps: 1
US-09-782-953-7 (1-198) x CGU60263 (1-2216)
QY 1 MetAspPheArgAspPheSerTyrAsnPheSerSerLeuIleAlaCysValAlaAsnAsp 20
DB 70 ATGCATTTTAGGACCTTTAACTACAAATTTTAGCTCCCTGATTCCTGTGGCAACCGGT 129
QY 21 AspValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40
DB 130 GATGTCCTCAGCGAAGTGAAACCGAGGCCAAATTTGAATCCCTCTTCAGGACGTATGAC 189
QY 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60
DB 190 AAGGACATCACCCTTCAGTATTTTAAGAGCTTCAACAGCTGTCGAATAAATTCAGCAAC 249
QY 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
DB 250 CCCTTATCCGAGCTGACGCCGCTGCGATGCGATGCGATGCGATGCGATGCGATGCG 309
QY 81 MetLysLeuTyrPheAlaGlnThrLeuHisLysGlySerSerHisLeuAlaProProAsn 100
DB 310 ATGAAGCTGTACTTCTCAGAGCTTTACATAGGAAGCTCACACCTCGCTCCGCCCAAT 369
QY 101 ProAspLysGlnPheLeuIleSerProProAlaSerProValGlyTyrLysGlnVal 120
DB 370 CCAGACAGCAGTTCCTCATCTCTCCCGGCTCTCCACCGAGTGGTTGGGAAGCAAGTA 429
QY 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro 140
DB 430 GAAGAGCTACCCAGTCTATAATTAAGATCTTTTATATGCTATCTCCAAGCTGGGCGCA 489
QY 141 GlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160
DB 490 GGTGAGAGTATGAGTGTGATGCGACCAACAGACACCACTCCACAGTGTGTAGTCCACGTG 549
QY 161 CysGluSerAspGlnGluAsnGluGluGluGluMetGluArgMetLysArgPro 180
DB 550 TGTGAAGAGCGCAAGAGAAAT---GAAGAGGAGGAGGAGATGAGAGAATGAAGAGACCC 606
QY 181 LysProLysIleIleGlnThrArgArgProGluTyrThrProIleHisLeuSer 198
DB 607 AAGCCAAAATTTATCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
RESULT 7
AX365321 2331 bp DNA linear PAT 15-FEB-2002
LOCUS AX365321
DEFINITION Sequence 11 from Patent WO0204491.
ACCESSION AX365321
VERSION AX365321.1 GI:18697049
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

AUTHORS Williams, S.R. and Rothermel, B.
TITLE Methods and compositions relating to muscle selective calcineurin
 interacting protein (mcip)
JOURNAL Patent: WO 0204491-A 11 17-JAN-2002;
 Board of Regents, The University of Texas System (US) ; Williams,
 Sanders R. (US) ; Rothermel, Beverly (US)
Location/Qualifiers

FEATURES
SOURCE

CDS

1..2331
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 144..737
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD23811.1"
 /db_xref="GI:18697050"
 /translation="MHFRNPNVSSSLIACVANSDFSESETRAKESLPRTYDKDIT
 FQYRSPKRVKINSPNSADARQLQHKTEFLKEMKLYPAQTLHGSSHLAPND
 KQFLISPPASPPVGMKVEDATPVINVDLVAISKPGKYEELHAAITDTPSVVHV
 CESDQKEEEMERMRPKKIOTRRPEYPIHLS"

BASE COUNT 630 a 470 c 547 g 684 t

ORIGIN

Alignment Scores:
 Pred. No.: 3,16e-88 Length: 2331
 Score: 970.50 Matches: 185
 Percent Similarity: 96.97% Conservative: 7
 Best Local Similarity: 93.43% Mismatches: 5
 Query Match: 93.05% Indels: 1
 DB: 6 Gaps: 1

US-09-782-953-7 (1-198) x AX365321 (1-2331)

QY 1 MetAapPhaArgAapPheSerTyraPheSerSerLeuIleAlaCyValAlaAenAap 20
 DB 144 ATGCATTTAAGAACTTTAACTACAGTTTACCTCTGATTCCTGCTGGCAACACT 203
 QY 21 AspValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThrTyraP 40
 DB 204 GATATCTTCGCAAGTAAAGTAAACAGGGCCAAATTGAGTCCCTCTTAGAGAGTATGAC 263
 QY 41 LysAapThrThrPheGlnTyrrPheLysSerPheLysArgValArgIleAenPheSerAap 60
 DB 264 AAGACATCACCTTTCAGTATTTTAAGAGCTCAACAGAGTCAAGATTAACCTTCAGCAAC 323
 QY 61 ProLeuSerAlaAlaAapAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLeuGlu 80
 DB 324 CCTTCTCCGACAGATGCCAGCTCCAGCTGCATTAAGCTAGTTTCTGGAAAGGAA 383
 QY 81 MetLysLeuTyrrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProPheAap 100
 DB 384 ATGAGTTATATTTTGGCTGACACTTACATAGGAAGCTCACACCTGGCTCCGCCAAAT 443
 QY 101 ProAapLysGlnPheLeuIleSerProProAlaSerProProValGlyTrpLysGlnVal 120
 DB 444 CCGAGACAAAGAGTTTTCGATCTCCCTCCGCTCCGCGAGGGATGGAACAAAGTG 503
 QY 121 GluAapAlaThrProValIleAenTyrrAapLeuLeuTyrrAlaIleSerLysLeuGlyPro 140
 DB 504 GAAGATGCAAGCCAGTCATTAACATGATCTCTTAATGCCATCTCCAAAGCTGGGGCCA 563
 QY 141 GlyGluLysTyrrGluLeuHisAlaAlaThrAapProThrProSerValValValHisVal 160
 DB 564 GGGGAAAGATGAATTGCAAGCAGCAGTGCACCACTCCAGCGGTGGTCCATGTA 623
 QY 161 CysGluSerAapGlnGluAenGluGluGluGluGluMetGluAargMetLysAapArgPro 180
 DB 624 TGTGAGAGTGTCAAGAG---AAGAGAGAAAGAAAGAAATGGAAATAAGAGAGACCT 680
 QY 181 LysProLysIleIleGlnThrArgArgProGluTyrrThrProIleHisLeuSer 198
 DB 681 AAGCCAAATTAATTCAGACCAAGAGCCGAGAGTACAGCGGATCCACCTCAGC 734

RESULT 8

HSU85267
 LOCUS 2346 bp mRNA linear PRI 19-APR-2000
 DEFINITION Homo sapiens down syndrome candidate region 1 (DSCR1) gene,
 alternative exon 1, complete cds.
 ACCESSION U85267
 VERSION U85267.2 GI:7596915
 KEYWORDS

SOURCE

ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 2346)
 Fuenfies,J.J., Pritchard,M.A., Planas,A.M., Bosch,A., Ferrer,I. and
 Batiavill,X.
 A new human gene from the Down syndrome critical region encodes a
 proline-rich protein highly expressed in fetal brain and heart
 Hum. Mol. Genet. 4 (10), 1935-1944 (1995)
 96121593
 8595418

REFERENCE

JOURNAL
 MEDLINE
 PUBMED
 8595418

2 (bases 1 to 2346)
 Fuenfies,J.J., Pritchard,M.A. and Batiavill,X.
 Genomic organization, alternative splicing, and expression patterns
 of the DSCR1 (Down syndrome candidate region 1) gene
 Genomics 44 (3), 358-361 (1997)
 97468152

3 (bases 1 to 2346)
 Fuenfies,J.J., Pritchard,M. and Batiavill,X.
 Direct Submission
 Submitted (14-APR-1997) Genetica Molecular, Institut de Recerca
 Oncologica, Autovia de Castelldefels, km 2.7, Hospitalat, Barcelona
 08907, Spain
 4 (bases 1 to 2346)
 Fuenfies,J.J., Pritchard,M. and Batiavill,X.
 Direct Submission
 Submitted (19-APR-2000) Genetica Molecular, Institut de Recerca
 Oncologica, Autovia de Castelldefels, km 2.7, Hospitalat, Barcelona
 08907, Spain

REFERENCE
 AUTHORS
 TITLE
 JOURNAL

REMARK
 COMMENT
 FEATURES
 SOURCE

1..2346
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="21"
 /map="21q22.1-q22.2"
 1..2346
 /gene="DSCR1"
 144..737
 /gene="DSCR1"
 /note="down syndrome candidate region 1; one of four
 alternatively spliced exon 1"
 /codon_start=1
 /protein_id="AA894372.1"
 /db_xref="GI:2612868"

BASE COUNT

645 a 470 c 547 g 684 t

ORIGIN

Alignment Scores:
 Pred. No.: 3,16e-88 Length: 2346
 Score: 970.50 Matches: 185
 Percent Similarity: 96.97% Conservative: 7
 Best Local Similarity: 93.43% Mismatches: 5
 Query Match: 93.05% Indels: 1
 DB: 9 Gaps: 1

US-09-782-953-7 (1-198) x HSU85267 (1-2346)

QY 1 MetAapPhaArgAapPheSerTyrrAapPheSerSerLeuIleAlaCyValAlaAenAap 20

ORIGIN

Alignment Scores: 6,596-84 Length: 597
 Pred. No.: 920.50 Matches: 178
 Score: 920.42% Conservative: 5
 Percent Similarity: 92.42% Mismatches: 12
 Best Local Similarity: 89.90% Indels: 3
 Query Match: 88.26% Gaps: 1
 DB: 6

US-09-782-953-7 (1-198) x AX365312 (1-597)

Qy 1 MetAapPheArGaSPheSerTyRanPheSerLeuIleAlaCyValAlaAsnAp 20
 Db 10 GTGAGTCTGCAGGACCTGCGG-----AGCGCACCATGCGCTGCCACCTGACCGG 60
 Qy 21 AapValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThTyraP 40
 Db 61 CGCGTGTTCGTGACGCGCTGTGCGGCGCAAAATTGAAATCCCTTTCCAGAACATATGAC 120
 Qy 41 LyAapPheThThPheGlnTyRPhelYSerPheLysArgValArgIleAsnPheSerAen 60
 Db 121 AAGGACACCACTTCACAGATTTTAAGACTTCAACGTCGCGATTAACCTTCAGCAAC 180
 Qy 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
 Db 181 CCTTATCTGCAGCGCATGCCAGGCTGCGGCTGCACAAAGCGAGTTCCTGGGGAAGGAA 240
 Qy 81 MetLysLeuTyRPhelArgInThrLeuHisIleGlySerSerHisLeuAlaProProAen 100
 Db 241 ATGAAGTTGATTTGCTCAGACTTACATAGAAAGTTCAACCTGCTCCGCCCAAT 300
 Qy 101 ProAapLysGlnPheLeuLysSerProProAlaSerProProValGlyTyRPhyGlnVal 120
 Db 301 CCGGACAAACAGTTCCTCATCTCCCTCGGCTCTCCCTCGGTGGCTGGAAACAAGTA 360
 Qy 121 GluAapAlaThProValIleAenTyRaspLeuLeuTyRAlaIleSerLysLeuGlyPro 140
 Db 361 GAAGATGCCACCCCGCTCAATTAATAGATCTTTTATGCTTATCCATCCCAAGCTGGGCGCA 420
 Qy 141 GlyLysTyRGlulLeuHisAlaIaThrAapProThProSerAlaValAlaHisVal 160
 Db 421 GGAGAAAGATGATGAATGATGATGACGACGACACCCCACTCCAGTGTGTGTCACGAG 480
 Qy 161 CysGluSerAapGlnGluAenGluGluGluGluGluGluMetGluArgMetLysArgPro 180
 Db 481 TGTGAGATGACCAAGAAATGAGAGAAAGAAAGAAAGATGAGATGAAGAAGACCC 540
 Qy 181 LysProLysIleIleGlnThrArgArgProGluTyRThProIleHisLeuSer 198
 Db 541 AAGCCCAAAATCATCCAGACACGAGACCGAGATACACCGATCCACCTTACG 594

RESULT 11

AF237789 597 bp mRNA linear ROD 12-APR-2000
 LOCUS AF237789
 DEFINITION Mus musculus myocyte-enriched calineurin interactin protein 1
 splice variant 1 mRNA, complete cds.
 VERSION AF237789.1 GI:7542525
 KEYWORDS
 SOURCE Mus musculus
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;
 1 (bases 1 to 597)
 Williams, B.A., Vega, R.B., Yang, J., Wu, H., Basseel-Duby, R.S. and
 TITLE A Protein Encoded within the Down Syndrome Critical Region is
 Enriched in Striated Muscles and Inhibits Calcineurin Signaling
 J. Biol. Chem. (2000) In press
 REFERENCE 2 (bases 1 to 597)
 Rothermel, B.A., Vega, R.B., Yang, J., Wu, H., Basseel-Duby, R.S. and
 AUTHORS Williams, R.S.

TITLE Direct Submission
 JOURNAL Submitted (23-FEB-2000) Internal Medicine, University of Texas
 Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, TX
 75390, USA

FEATURES

source location/Qualifiers
 1..597
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /chromosome="16"
 1..597
 /note="similar to Homo sapiens DSCR1; MCFP1; similar to
 EST sequence of Genbank Accession Number AA734360"

CDS
 /product="myocyte-enriched calcineurin interactin protein
 1 splice variant 1"
 /protein_id="AA63485.1"
 /db_xref="GI:7542526"
 /translation="MEEVDLQDLSATLACHLDPRVVDGLCRAPESLFRDYDKRT
 FQYFSKRRIRINFSNPLSADRLRLHRTKTEFKEMKLYFAOTLHIGSSHLAPNPD
 KQFLISPPASPVGMKVEDATPVINDLVAISKGPGEKELHAAATDPTPSVYVAV
 CESDQENESEEEEMERMKRPKRIQTRREPTPIHLS"

BASE COUNT 160 a 170 c 151 g 116 t

ORIGIN

Alignment Scores: 6,596-84 Length: 597
 Pred. No.: 920.50 Matches: 178
 Score: 920.42% Conservative: 5
 Percent Similarity: 92.42% Mismatches: 12
 Best Local Similarity: 89.90% Indels: 3
 Query Match: 88.26% Gaps: 1
 DB: 10

US-09-782-953-7 (1-198) x AE237789 (1-597)

Qy 1 MetAapPheArGaSPheSerTyRanPheSerLeuIleAlaCyValAlaAsnAp 20
 Db 10 GTGAGTCTGCAGGACCTGCGG-----AGCGCACCATGCGCTGCCACCTGACCGG 60
 Qy 21 AapValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThTyraP 40
 Db 61 CGCGTGTTCGTGACGCGCTGTGCGGCGCAAAATTGAAATCCCTTTCCAGAACATATGAC 120
 Qy 41 LyAapPheThThPheGlnTyRPhelYSerPheLysArgValArgIleAsnPheSerAen 60
 Db 121 AAGGACACCACTTCACAGATTTTAAGACTTCAACGTCGCGATTAACCTTCAGCAAC 180
 Qy 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
 Db 181 CCTTATCTGCAGCGCATGCCAGGCTGCGGCTGCACAAAGCGAGTTCCTGGGGAAGGAA 240
 Qy 81 MetLysLeuTyRPhelArgInThrLeuHisIleGlySerSerHisLeuAlaProProAen 100
 Db 241 ATGAAGTTGATTTGCTCAGACTTACATAGAAAGTTCAACCTGCTCCGCCCAAT 300
 Qy 101 ProAapLysGlnPheLeuLysSerProProAlaSerProProValGlyTyRPhyGlnVal 120
 Db 301 CCGGACAAACAGTTCCTCATCTCCCTCGGCTCTCCCTCGGTGGCTGGAAACAAGTA 360
 Qy 121 GluAapAlaThProValIleAenTyRaspLeuLeuTyRAlaIleSerLysLeuGlyPro 140
 Db 361 GAAGATGCCACCCCGCTCAATTAATAGATCTTTTATGCTTATCCATCCCAAGCTGGGCGCA 420
 Qy 141 GlyLysTyRGlulLeuHisAlaIaThrAapProThProSerAlaValAlaHisVal 160
 Db 421 GGAGAAAGATGATGAATGATGATGACGACGACACCCCACTCCAGTGTGTGTCACGAG 480
 Qy 161 CysGluSerAapGlnGluAenGluGluGluGluGluGluGluMetGluArgMetLysArgPro 180
 Db 481 TGTGAGATGACCAAGAAATGAGAGAAAGAAAGAAAGATGAGATGAAGAAGACCC 540
 Qy 181 LysProLysIleIleGlnThrArgArgProGluTyRThProIleHisLeuSer 198
 Db 541 AAGCCCAAAATCATCCAGACACGAGACCGGAGATACACCGATCCACCTTACG 594


```

RESULT 12
AF282255
LOCUS
DEFINITION
Mus musculus Down syndrome candidate region 1 protein (Dscr1) mRNA, complete cds.
ACCESSION
AF282255
VERSION
AF282255.1 GI:9652249
KEYWORDS
Mus musculus.
SOURCE
Mus musculus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 621)
Strippoli, P., Petrini, M., Lenzi, L., Carinci, P. and Zannotti, M. The murine DSCR1-like (Down syndrome candidate region 1) gene family: conserved synteny with the human orthologous genes
JOURNAL
Gene 257 (2), 223-232 (2000)
MEDLINE
20534792
PUBMED
11080588
REFERENCE
2 (bases 1 to 621)
Strippoli, P., Petrini, M., Lenzi, L., Carinci, P. and Zannotti, M. Direct Submission
AUTHORS
Submitted (24-JUN-2000) Istituto di Istologia ed Embriologia Generale, Universite di Bologna, Via Belmeloro, 8, Bologna, BO 40126, Italy
JOURNAL
Location/Qualifiers
FEATURES
source
1..621
/organism="Mus musculus"
/strain="BALB/c"
/db_xref="taxon:10090"
/chromosome="16"
/sex="male"
/tissue_type="brain"
1..621
/gene="Dscr1"
8..604
/gene="Dscr1"
/note="similar to the Homo sapiens Down syndrome candidate region 1 protein"
/codon_start=1
/product="Down syndrome candidate region 1 protein"
/protein_id="AAF91461.1"
/db_xref="GI:9652250"
/tranlation="MEEVDLQIPSAIACHLDRVFDGLCRAPESLFRVYDKDTT FQPKFKRVINFPISAAARLHKTEFLGKMKLYFAQTLHGSSHLAPNPD KQFLIPSPAPVPGWQVEDATPVINLLYAIKLGPGKEKVELHAATDTPTSVVHV QESDQNEEEEEEMERKPKPKIIQTRPEVTPHLS"
BASE COUNT 166 a 179 c 159 g 117 t
ORIGIN

Alignment Scores:
Pred. No.: 4.43e-83 Length: 621
Score: 912.50 Matches: 177
Percent Similarity: 91.92% Conservative: 5
Best Local Similarity: 89.39% Mismatches: 13
Query Match: 87.49% Indels: 3
DB: 10 Gaps: 1

US-09-782-953-7 (1-198) x AF282255 (1-621)

Qy 1 MetAspPheArgAspPheSerTyrAsnPheSerLeuLeuAlaCysValAlaAsnAsp 20
Db 17 GTGGATCTGCAGGACCTGCCG-----AGCGCCACCATCGCTGCCACCTGGACCCG 67

Qy 21 AspValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40
Db 68 CGCGTGTTCTGGACGCGCTGTGCCGCGCCAAATTTGAATCCCTCTTCAGAAATATGAC 127

Qy 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgLeuAsnPheSerAsn 60
Db 128 AAGGACACCACTTCAGTATTTAAAGAGCTTCAACAGTGTCCGAGATAAACTTCAGCAAC 187

```

```

Qy 61 ProLeuSerAlaAlaAspAlaArgLeuAtcLeuHisLysThrGluPheLeuGlyLysGlu 80
Db 188 CCTTATCTGAGCCGATGCCAGCTCGGCTGCACAAAGACCAGATTCTCTGGGGAAGGAA 247

Qy 81 MetLysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProProAsn 100
Db 248 ATGAAGTTGTATTGTCTCAGACTTTTACACATAGGAAGTTTCACACCTGGCTCCGCCAAT 307

Qy 101 ProAspLysGlnPheLeuIleSerProProAlaSerProProValGlyTyrLysGlnVal 120
Db 308 CCCGACAAACAGTTCTCATCTCCCTCCGGCCTCTCTCCCTGGCTGGAACAAGTA 367

Qy 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro 140
Db 368 GAAGATGCCACCCCGCTCATAAATTACGATCTTTTATATGCCATCTCCCAAGCTGGGGCCA 427

Qy 141 GlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160
Db 428 GGAGAAAGTATGAACCTGCATGCAGCGACAGACACCACTCCCACTGTGTGTCCAGTG 487

Qy 161 CysGluSerAspGlnGluGluGluGluGluMetGluArgMetLysArgPro 180
Db 488 TGTGAGAGTGCCCAAGAGATGAGGAGGAGGAGGAGATGGAGAGATGAAGAGACCC 547

Qy 181 LysProLysIleIleGlnThrArgArgProGluTyrThrProIleHisLeuSer 198
Db 548 AAGCCCAAAATCATCCAGACACGAGACCGAGTACACACCCATCCCTCAGC 601

RESULT 13
AF260717
LOCUS
DEFINITION
Mus musculus Down syndrome candidate region 1 (Dscr1) mRNA, complete cds.
ACCESSION
AF260717
VERSION
AF260717.1 GI:7839596
KEYWORDS
Mus musculus.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 2125)
Casas, C., Martinez, S., Pritchard, M.A., Fuentes, J.J., Nadal, M., Guimera, J., Arbones, M., Florez, J., Soriano, E., Estivill, X. and Alcantara, S.
Dscr1, a novel endogenous inhibitor of calcineurin signaling, is expressed in the primitive ventricle of the heart and during neurogenesis
JOURNAL
Mech. Dev. 101 (1-2), 289-292 (2001)
MEDLINE
21152920
PUBMED
11231093
REFERENCE
2 (bases 1 to 2125)
Fuentes, J.J., Pritchard, M., Pucharcas, C. and Estivill, X. Down syndrome candidate region 1 (Dscr1), one of three alternatively spliced exon 1
AUTHORS
Unpublished
JOURNAL
3 (bases 1 to 2125)
Fuentes, J.J., Pritchard, M., Pucharcas, C. and Estivill, X. Direct Submission
AUTHORS
Submitted (26-APR-2000) Centre de Genetica Medica i Molecular, IRO, Avia. Cascelldefels Km. 2.7, L'Hospitalet de Llobregat, Barcelona 08907, Spain
FEATURES
Location/Qualifiers
source
1..2125
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="16"
1..2125
/gene="Dscr1"
15..611
/gene="Dscr1"
/note="calcineurin inhibitor"
/codon_start=1
/product="Down syndrome candidate region 1"

```



```

RESULT 15
AX365324
LOCUS      AX365324
DEFINITION Sequence 14 from Patent WO0204491.
ACCESSION AX365324
VERSION    AX365324.1 GI:18697051
KEYWORDS
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1
AUTHORS    Williams, S.R. and Rothermel, B.
TITLE      Methods and compositions relating to muscle selective calcineurin
            interacting protein (ncip)
JOURNAL    Patent: WO 0204491-A 14 17-JAN-2002;
            Board of Regents, The University of Texas System (US); Williams,
            Sanders R. (US); Rothermel, Beverly (US)
FEATURES   Location/Qualifiers
            1..2212
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
             25..618
              /note="unnamed protein product"
              /codon_start=1
              /protein_id="CAD23812.1"
              /db_xref="GI:18697052"
              /translation="MEEVDLQDLPSTAIACHLDPRVFVDGLCRKFESLFRVYDKDIT
              FQYFKFRVIRNFPSAADARLQHLKTEFLGKEMKLYPAQTLHGSSHLAPNPD
              KQFLISPPASPPVQWQVATPVINVDLLYAIKLGPGKEKYLHAATTTTSVVVHV
              CESDQKEEEMRRPKPIQTRRPEYPIHLS"
BASE COUNT 588 a 455 c 529 g 640 t
ORIGIN
Alignment Scores:
Pred. No.:      1.26e-78      Length:      2212
Score:          875.00      Matches:      171
Percent Similarity: 90.40%      Conservative: 8
Best Local Similarity: 86.36%      Mismatches: 15
Query Match:     83.89%      Indels:      4
DB:              6           Gaps:      2
US-09-782-953-7 (1-198) x AX365324 (1-2212)
QY  1 MetAspPheArgAspPheSerTyrAsnPheSerSerLeuIleAlaCysValAlaAsnAsp 20
Db  34 GTGGACCTGCAGGACCTGCCCC-----AGCGCCACCACCTGCTGCACCTGGACCCG 84
QY  21 AspValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40
Db  85 CGCGTGTTCGTGGACGGCCCTGTCCCGGCCCAATTGAGTCCCTCTTTAGGACGTATGAC 144
QY  41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60
Db  145 AAGGACATCACCTTTTCAGTATTTTAAAGAGCTTCAACAGCAGTCAGATAAATTCAGCAAC 204
QY  61 ProLeuSerAlaAlaPheAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
Db  205 CCCTTCTCCGACGAGATGCCAGGCTCCAGCTGCATAAGACTGAGTTTCTGGGAAAGGAA 264
QY  81 MetLysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProProAsn 100
Db  265 ATGAAGTTATATTTTCTCAGACCTTACACATAGGAGCTCACACCTGGCTCCGCCAAAT 324
QY  101 ProAspLysGlnPheLeuIleSerProProAlaSerProProValGlyTyrLysGlnVal 120
Db  325 CCAGACAAGCAGATTCTGATCTCCCTCCCGCTCTCCGCGAGTGGATGGAAACAAAGTG 384
QY  121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro 140
Db  385 GAAGATGCGGACCCAGTCATAAACTATGATCTCTTATATGCCATCTCCAAAGCTGGGCCA 444
QY  141 GlyLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160

```

```

Db  445 GGGGAAAAAGTATGAATTGCACGCGAGCTGCACACCACTCCCGAGCTGGTGGTCCATGTA 504
QY  161 CysGluSerAspGlnGluAsnGluGluGluGluMetGluArgMetLysArgPro 180
Db  505 TGTGAGAGTGTATCAAGAG---AAGGAGGAGAAAGAGAGAAATGGAAGAATGAGGAGACCT 561
QY  181 LysProLysIleIleGlnThrArgArgProGluTyrThrProIleHisLeuSer 198
Db  562 AAGCCAAAATATTCAGACCAAGAGGCGGAGGTACACGCCGATCCACCTCAGC 615

```

Search completed: December 14, 2002, 21:32:04
Job time : 1661.79 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 14, 2002, 19:56:01 ; Search time 1118.55 Seconds
(without alignments)
2866.843 Million cell updates/sec

Title: US-09-782-953-7
Perfect score: 1043
Sequence: 1 MDRFDFSYNFSLIACVAND.....RPKPKIQTTRPEYTHLS 198

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DBV=xlh
-Q=/cgn2_1/USPTO.spool/US09782953/runat_11122002_114429_17353/app.query.fasta_1.1173
-DB=EST -QFMT=fastap -SUFFIX=txt -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bites -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pcr -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09782953 -CGN_1_1583 -RUNAT_11122002_114429_17353 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAPP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST.*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_htc.*
9: gb_est1.*
10: gb_est2.*
11: gb_htc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: gb_gas.*
18: em_gas_hum.*
19: em_gas_inv.*
20: em_gas_pin.*
21: em_gas_vrt.*
22: em_gas_fun.*
23: em_gas_nam.*
24: em_gas_mus.*
25: em_gas_other.*
26: em_gas_pro.*
27: em_gas_rod.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1021	97.9	1158	14	BQ895506	BQ895506 AGENCOURT
2	974	93.4	753	13	B1148584	B1148584 602911995
3	970.5	93.0	837	9	AU124628	AU124628 AU124628
4	970.5	93.0	931	9	AL543576	AL543576 AU543576
5	970.5	93.0	939	9	AL546617	AL546617 AU546617
6	970.5	93.0	1041	13	BM450020	BM450020 AGENCOURT
7	963.5	92.4	946	14	BQ278576	BQ278576 AGENCOURT
8	961.5	92.2	718	9	AU131040	AU131040 AU131040
9	957.5	91.8	885	9	AL544755	AL544755 AU544755
10	941.5	90.3	662	13	BI464521	BI464521 603206117
11	938.5	90.0	939	9	AL554686	AL554686 AU554686
12	937	89.8	710	10	BB617325	BB617325 BB617325
13	934.5	89.6	906	9	AL544313	AL544313 AU544313
14	932.5	89.4	1078	13	BM541636	BM541636 AGENCOURT
15	929.5	89.1	740	13	BI463566	BI463566 603206210
16	925	88.7	950	12	BF782633	BF782633 603107462
17	924	88.6	665	13	BI219142	BI219142 602935788
18	909.5	87.2	811	12	BG570239	BG570239 602590641
19	906.5	86.9	769	13	BI767955	BI767955 603061320
20	904.5	86.7	827	13	BI102432	BI102432 602889619
21	898.5	86.1	980	12	BG296537	BG296537 602394441
22	896	85.9	711	13	BG915568	BG915568 602815815
23	887.5	85.1	710	12	BG475986	BG475986 602521111
24	885.5	84.9	1507	11	AK010696	AK010696 Mus muscu
25	875	83.9	828	9	AL536447	AL536447 AL536447
26	875	83.9	890	9	AL538796	AL538796 AL538796
27	868	83.2	599	9	AA200984	AA200984 mu12e11.r
28	862	82.6	867	14	BQ427531	BQ427531 AGENCOURT
29	861.5	82.6	751	9	AL576189	AL576189 AL576189
30	850.5	81.5	1014	9	AL559594	AL559594 AL559594
31	848.5	81.4	689	9	AL652662	AL652662 AL652662
32	847	81.2	937	12	BE795722	BE795722 601590730
33	846	81.1	956	9	AL556803	AL556803 AL556803
34	842.5	80.8	875	12	BG574693	BG574693 602596709
35	840.5	80.6	642	14	BQ520533	BQ520533 NISC_1105
36	833.5	79.9	898	14	BQ733563	BQ733563 AGENCOURT
37	829.5	79.5	1047	12	BG287042	BG287042 602382338
38	829	79.5	923	9	AL551657	AL551657 AL551657
39	823.5	79.0	664	13	BJ061844	BJ061844 603385358
40	815.5	78.2	627	13	BI859506	BI859506 603385358
41	808	77.5	520	14	BQ749142	BQ749142 UI-M-FD0
42	793.5	76.1	822	12	BE895008	BE895008 601433720
43	791.5	75.9	986	12	BG828534	BG828534 602752653
44	784.5	75.2	599	13	BJ064020	BJ064020 BJ064020
45	779.5	74.7	596	13	BJ058360	BJ058360 BJ058360

ALIGNMENTS

RESULT 1
BQ895506
LOCUS
DEFINITION BQ895506 8762889 NIH_MGC_130 Mus musculus cDNA clone IMAGE:6329174
5', mRNA sequence.
ACCESSION BQ895506
VERSION BQ895506.1 GI:22287520
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1156)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)


```

QY 81 MetLysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProProAsn 100
Db 328 ATGAAGTG-TATTTGCTCAGACTTTACATAGGAAGTTCACACCTGGCTCCGCCCAAT 386

QY 101 ProAspLysGlnPheLeuSerProProAlaSerProProValGlyTyrPheGlnVal 120
Db 387 CCGGCAACAGTTCCTCATCTCCCTCCGCCCTCTCTCCCTGGCTGGGAAACAAGTA 446

QY 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro 140
Db 447 GAAGATGCCACCCCGTCGTCATAAATACGATCTTTATATGTCATCTCAAGCTGGGCCA 506

QY 141 GlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160
Db 507 GGAGAGAAAGTATGAATGTCATGTCGACGACACACCACTCCAGTGTGGTCCACGTTG 566

QY 161 CysGluSerAspGlnGluAsnGluGluGluGluGluMetGluArgMetLysArgPro 180
Db 567 TGTGAGAGTGACCAAGAGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 626

QY 181 LysProLysIleIleGlnThrArgProGluTyrThrProIleHisLeuSer 198
Db 627 AAGCCAAAATCATCCAGACGCGGAGCGGAGTACACACCCATCCACCTCAGC 680

RESULT 3
LOCUS AUI24628 837 bp mRNA linear EST 01-AUG-2002
DEFINITION AUI24628 NT2RM4 Homo sapiens cDNA clone NT2RM4000307 5', mRNA
sequence.
ACCESSION AUI24628
VERSION AUI24628.1 GI:10949344
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Ota,T., Wakamatsu,A., Ozawa,M., Ishii,S., Saito,K., Yamamoto,J.,
Nakamura,Y., Nishikawa,T., Nagai,T., Suzuki,Y., Sugano,S. and
Ishigai,T.
TITLE HRI human cDNA project (Ota,T., Wakamatsu,A., Ozawa,M., Ishii,S.,
Saito,K., Yamamoto,J., Nakamura,Y., Nishikawa,T., Nagai,T., Suzuki
Y., Sugano,S., Ishigai,T.)
JOURNAL Unpublished (2000)
COMMENT Contact: Takao Ishigai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing; Helix
Research Institute; cDNA library construction; Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
FEATURES
Location/Qualifiers
source
1..837
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NT2RM4000307"
/cell_type="teratocarcinoma"
/cell_line="NT2"
/Note="Vector: pMB18FL3; mRNA from uninduced NT2 neuronal
precursor cells"
BASE COUNT 241 a 201 c 207 g 186 t 2 others
ORIGIN
Alignment Scores:
Pred. No.: 1.55e-98 Length: 837
Score: 970.50 Matches: 185
Percent Similarity: 96.97% Conservative: 7

```

```

Best Local Similarity: 93.43% Mismatches: 5
Query Match: 93.05% Indels: 1
DB: 9 Gaps: 1
US-09-782-953-7 (1-198) x AUI24628 (1-837)

QY 1 MetAspPheArgAspPheSerTyrAsnPheSerSerLeuIleAlaCysValAlaAsnAsp 20
Db 118 ATGCATTTTAGAAACCTTTAACTACAGCTTTAGCTCCCTGATTCCTGCTGTGGCAACAGT 177

QY 21 AspValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40
Db 178 GATATCTTCAGCGAAGTGAACACCGGCCAAATTTAGTCCCTCTTTAGGACGATATGAC 237

QY 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60
Db 238 AAGGACATCACCTTTTCAGTATTTTAAGAGCTTCAAACAGAGTCAGAATAAATTCAGCAAC 297

QY 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
Db 298 CCTTCTCCGACGACAGATGCCAGGCTCCAGCTGCATAAGACTGAGTTCTGGGAAGGAA 357

QY 81 MetLysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProProAsn 100
Db 358 ATGAAGTTATATTTTGCTCAGACCTTACACATAGGAGCTCACACTGGCTCCGCCAAT 417

QY 101 ProAspLysGlnPheLeuSerProProAlaSerProProValGlyTyrPheGlnVal 120
Db 418 CCAGCAACAGCAGTCTTCATCTCCCTCCGCCCTCCGCCAGTGGATGGAACAAGTG 477

QY 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro 140
Db 478 GAAGATGCCACCCAGTCATAAATATATCTTTATATGCAATCTCCAAGCTGGGGCA 537

QY 141 GlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160
Db 538 GGGGAAAAGTAGTAATTCATGCATGCAGCAGTACACCACTCCAGCGCTGGTCCATGTA 597

QY 161 CysGluSerAspGlnGluAsnGluGluGluGluGluMetGluArgMetLysArgPro 180
Db 598 TGTGAGAGTGATCAAGAG---AAGGAGGAGGAAGGAAATGGAAGAATGAGGAGACCT 654

QY 181 LysProLysIleIleGlnThrArgArgProGluTyrThrProIleHisLeuSer 198
Db 655 AAGCCAAAATTCATCCAGACGCGGAGCGGAGTACACCGCGATCCACCTCAGC 708

RESULT 4
LOCUS AL543576 931 bp mRNA linear EST 16-FEB-2001
DEFINITION AL543576 LTI NFL006 PL2 Homo sapiens cDNA clone CS0DI006YK09 5
prime, mRNA sequence.
ACCESSION AL543576
VERSION AL543576.1 GI:12876055
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 931)
AUTHORS Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
source
1..931
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DI006YK09"
/clone_lib="LTI NFL006 PL2"
/tissue_type="placenta"

```

/note="Vector: pcwvSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pcwvSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 262 a 217 c 238 g 214 t
ORIGIN

Alignment Scores: 1.76e-98 Length: 931
Pred. No.: 970.50 Matches: 185
Score: 96.97% Conservative: 7
Percent Similarity: 93.43% Mismatches: 5
Best Local Similarity: 93.05% Indels: 1
Query Match: 9 Gaps: 1
DB:

US-09-782-953-7 (1-198) x AL54617 (1-931)

QY 1 MetAapPheArGApPheSerTyraPheSerSerLeuileAlaCyValAlaAsnAap 20
DB 50 ATGCATTTTGAAGAACTTAACTAAGTTTAACTCCCTGATGCTGCTGGCAACAGT 109
QY 21 AapValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThrTyraap 40
DB 110 GATATCTTCAGCAAACTGAAACCAAGGCCAAATTGAGTCCCTCTTAAAGACGATGAC 169
QY 41 LysAapThrThrPheGlnTyPheLysSerPheLysValArgileAsnPheserAan 60
DB 170 AAGGACATCCCTCTTCAATTTTAAAGACTTCAAGACTCAAGATTAACCTCAGCAAC 229
QY 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
DB 230 CCTCTTCCCGACAGATGCGCAGGCTCCAGCTGCATTAAGACTGATTTCTGGGAAAGGAA 289
QY 81 MetLysLeuTyPheAlaGlnThrLeuHisLysGlySerSerHisLeuAlaProProAan 100
DB 290 ATGAAGTAAATTTTGTCTGACACTTACACATGAGAAAGCTCACTGGCTCCGCCAAAT 349
QY 101 ProAapLysGlnPheLeuLysSerProAlaSerProProValGlyTyPlyGlnVal 120
DB 350 CCAACAAAGCAATTTCTGATCTCCCTCCGCTCCGCAAGTGGATGGAACAAAGTG 409
QY 121 GluAapAlaThrProValIleAsnTyraPheLeuTyraLisSerLysLeuGlyPro 140
DB 410 GAAGATCGAGCCCAAGTCATTAATGATCTTTATATGCAATCTCCAAAGCTGGGGCCA 469
QY 141 GlyLysLeuTyPheGluHisLysAlaLysThrAapProThrProSerValIleHisVal 160
DB 470 GGGGAAAGATGATGACGACGACGACGACCTCCCAAGCTGGTGGTCAAGTGA 529
QY 161 CysGluSerAapGlnGluAenGluGluGluGluMetGluArgMetLysArgPro 180
DB 530 TGTGAGACTGATCAAGG---AAGAGAGAAAGAGGAAATGAAAGATAGAGAGACT 586
QY 181 LysProLysIleIleGlnThrArgArgProGluTyThrProIleHisLeuSer 198
DB 587 AAGCAAAAATTATCCAGACGAGGCCAGAGTACACGCGGATCCACTGAGC 640

RESULT 5
AL54617 AL546617 939 bp mRNA linear EST 16-FEB-2001
LOCUS AL546617 LTI_NFL006.PL2 Homo sapiens cDNA clone CS001029YJ21.5
DEFINITION prime, mRNA sequence.
ACCESSION AL546617
VERSION AL546617.1 GI:12879906
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
ATTNORS
TITLE
JOURNAL
COMMENT
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 939)
Li, W.B., Gruber, C., Jessup, J., and Polayes, D. Full-length cDNA libraries and normalization Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage BP 191 91006 Evry cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1. .939
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS01029YJ21"
/clone_lib="LTI_NFL006.PL2"
/tissue_type="placenta"
/note="Vector: pcwvSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pcwvSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 263 a 219 c 238 g 218 t
ORIGIN

Alignment Scores: 1.77e-98 Length: 939
Pred. No.: 970.50 Matches: 185
Score: 96.97% Conservative: 7
Percent Similarity: 93.43% Mismatches: 5
Best Local Similarity: 93.05% Indels: 1
Query Match: 9 Gaps: 1
DB:

US-09-782-953-7 (1-198) x AL54617 (1-939)

QY 1 MetAapPheArGApPheSerTyraPheSerSerLeuileAlaCyValAlaAsnAap 20
DB 53 ATGCATTTTGAAGAACTTAACTAAGTTTAACTCCCTGATGCTGCTGGCAACAGT 112
QY 21 AapValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThrTyraap 40
DB 113 GATATCTTCAGCAAACTGAAACCAAGGCCAAATTGAGTCCCTCTTAAAGACGATGAC 172
QY 41 LysAapThrThrPheGlnTyPheLysSerPheLysValArgileAsnPheserAan 60
DB 173 AAGGACATACCTTTCAATTTTAAAGCTTCAACAGACTCAAGATTAACCTCAGCAAC 232
QY 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
DB 233 CCTCTCCGACAGATGCCAGCTCCAGCTGCAATGAGACTGATTTCTGGGAAAGGAA 292
QY 81 MetLysLeuTyPheAlaGlnThrLeuHisLysGlySerSerHisLeuAlaProProAan 100
DB 293 ATGAAGTAAATTTTGTCTGACACTTACACATGAGAAAGCTCACTGGCTCCGCCAAAT 352
QY 101 ProAapLysGlnPheLeuLysSerProAlaSerProProValGlyTyPlyGlnVal 120
DB 353 CCAAGCAAGCAATTTTGTATCTCCCTCCGCTCCGCAAGTGGATGGAACAAAGTG 412
QY 121 GluAapAlaThrProValIleAsnTyraPheLeuTyraLisSerLysLeuGlyPro 140
DB 413 GAAGATCGAGCCCAAGTCATTAATGATCTTTATATGCAATCTCCAAAGCTGGGGCCA 472
QY 141 GlyLysLeuTyPheGluHisLysAlaLysThrAapProThrProSerValIleHisVal 160
DB 473 GGGGAAAGATGATGACGACGACGACGACCTCCCAAGCTGGTGGTCAAGTGA 532
QY 161 CysGluSerAapGlnGluAenGluGluGluGluMetGluArgMetLysArgPro 180

Qy 181 LysProLysIlelleGlnThrArgArgProGluTyrThrProIleHisLeuSer 198
 Db 639 AAGCAAAATTTATCCAGACCGGANGCGGAGTACACGCCGATCCACTTAAAC 692

RESULT 9
 AL544755
 LOCUS
 DEFINITION AL544755 LTI_NFL006_PL2 Homo sapiens cDNA linear EST 16-FEB-2001
 prime, mRNA sequence.
 AL544755
 ACCESSION AL544755.1 GI:12877235
 VERSION
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 885)
 Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
 source
 1..885
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS01022YJ23"
 /clone_lib="LTI_NFL006_PL2"
 /tissue_type="placenta"
 /notes="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
 Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com"

BASE COUNT 247 a 211 c 228 g 198 t 1 others
 ORIGIN

Alignment Scores:
 Pred. No.: 4.74e-97 Length: 885
 Score: 957.50 Matches: 195
 Percent Similarity: 96.48% Conservations: 7
 Best Local Similarity: 92.96% Mismatches: 5
 Query Match: 91.80% Indels: 2
 DB: 9 Gaps: 1

US-09-782-953-7 (1-198) x AL544755 (1-885)

Qy 1 MetAspPheArgAspPheSerTyrAsnPheSerLeuIleAlaCysValIleAsnAsp 20
 Db 50 ATGCATTTTAGAAACCTTTAACTACAGTTTAGTCCTGCTGTTGGCAACAGT 109

Qy 21 AspValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40
 Db 110 GATATCTTCAGCGAAGTAGAACACCGGCCAAATTTAGTGTCCTCTTTAGGACGATGAC 169

Qy 41 LysAspThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60
 Db 170 AAGGACATCACCTTTTCAGTATTTTAAGAGCTTCAAACGAGTCAGATAAATTCAGCAAC 229

Qy 61 ProLeuSerAlaAlaAspIleArgLeuArgLeuHisGlyThrGluPheLeuGlyLysGlu 80
 Db 230 CCCTTCCTCGAGCAGATCCAGGCTCCAGCTGCATAGAGCTCAGTTCTCGGAAGGAA 289

Qy 81 MetLysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProProAsn 100
 Db 290 ATGAAGTTATATTTTCTCAGACCTTACACATAGGAGCTCACACCTGCTCCGCAAT 349

21

was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371

Email : fliang@lifeotech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT	262 a	219 c	238 g	218 t	2 others
ORIGIN					
Alignment Scores:					
Pred. No.:	6.84e-95	Length:	939		
Score:	938.50	Matches:	183		
Percent Similarity:	95.00%	Conservative:	7		
Best Local Similarity:	91.50%	Mismatches:	7		
Query Match:	89.98%	Indels:	3		
DB:	9	Gaps:	1		

US-09-782-953-7 (1-198) x AL554686 (1-939)

Qy	1	MetaepphearAasppheserTyraaspheserleuilealaCyvaIaasnaasp	20
Db	53	ATGCATTTTAAGAACTTAACACTACAGTTTAGCTCCCGATGGCTGTGTGGCAACAGT	112
Qy	21	AapValIpheserGJuserGJuthrarg--AAlaypHeGJuserleuDeuPheargrThrya	40
Db	113	GATATCTTCAGCAAGAGCAACCATGTGGCCAAATTGAATGCCCTCTTTAGGACGTATG	172
Qy	40	aplyAasprThrPhheGlnTyrrPhelysersPhelysArxValArgIleasphesera	60
Db	173	ACAAGACATCCCTTCGATTTTAAGCTTCMAACGAGTCAGATTAACCTTCAGCA	232
Qy	60	anProleuserIaAlaaspaIargLeuArgLeuHisIelysThrgIubheugIylysg	80
Db	233	ACCCCTCTCCGCAACAGATGCCAGGCTCCAGCTGCATTAAGACTGATTCCTGGGAAAGG	292
Qy	80	IumetlyleuTyrrPhelaIagInTrLeuHisIleglyserSerHisIaleuIaProProa	100
Db	293	AAATGAAGTTATATTTTGTCTCAGACCTTAACATGTGAAGATCAACCTGGCTCCGCA	352
Qy	100	anProasprIyegInheleuIeserProProIaSerProProValGlyTPhlysgIny	120
Db	353	ATCCAGACAAAGCAATTCGATCTCCCTCCGCTCTCCGCCATGGGATGGAAACAG	412
Qy	120	aIglubspaIathrProValIleasenTyraaspleuTyrrAlaIeserlyleugIyp	140
Db	413	TGGAGATCCGACCCAGCATTAACATGATCTCTTATATCCATCTCCAAAGCTGGGGC	472
Qy	140	roglYglulYrTyrgIuleuHisIaIaIathrAapProThrProSerValValIhIay	160
Db	473	CAGGGGAAAGATGAATTGCACGCAAGCATCAACACCTCCGACGGGTGGTCCATG	532
Qy	160	aIcyGeluserAopngInuamGInuGInuGInuIumetGInuArymetlyBaryp	180
Db	533	TATGTAGAGTATCAAG--AAAGAGAGAAAGAGAAATGGAAAGATGAGGAC	589
Qy	180	rolyasProlysiIeIegInThrAryarProglUtyrThrProIleHisIeuser	198
Db	590	CTAAGCCAAAATTAATCCAGACCAAGAGCCGGAATACGCCGATCCACCTCAG	645

RESULT 12	
LOCUS	BB617325
DEFINITION	BB617325 710 bp mRNA linear EST 26-OCT-2001
ACCESSION	BB617325
VERSION	BB617325.1
KEYWORDS	EST.
SOURCE	house mouse
ORGANISM	Mus musculus
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;


```

Db 36% TCCAGACAGCAGTTCTGATCTCCCTCCCGCCTCTCCCGCAGTGGGATGGAAACAAGT 423
Qy 120 lGluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPr 140
Db 424 GGAAGATCGGACCCAGTCATAAACTATGATCTCTATATGCCATCTCCCAAGCTGGGGCC 483
Qy 140 oGlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVa 160
Db 484 AGGGAAAGATGATGAATTGCACGCGAGGACTGCACCACTCCCAAGCGTGGTGCCATGT 543
Qy 160 lCysGluSerAspGlnGluLeuGluGluGluGluMetGluArgMetLysArgPr 180
Db 544 ATGTGAGATGATCAAGAG---AAGGAGGAGAGAGAGAAATGGAAAGATGAGGAGCC 600
Qy 180 oLysProLysIlelleGlnThrArgArgProGluTyrThr-ProIleHisLeuSer 198
Db 601 TAAGCCAAAATTTATCCAGACAGGAGCGCGAGTACACGCCCGCATCCACTCAGC 656

RESULT 15
BI463566 740 bp mRNA linear EST 21-AUG-2001
LOCUS 603206210F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5271867 5',
DEFINITION mRNA sequence.
ACCESSION BI463566
VERSION BI463566.1 GI:15254222
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 740)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Tohiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1686 row: c column: 04
High quality sequence stop: 736.
FEATURES
Location/Qualifiers
1..740
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5271867"
/clone_lib="NIH_MGC_97"
/lab_host="DH10B"
/note="Organ: testis; Vector: pBluescriptR (modified
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTNN-3',
size-selected for average insert size 2.2 kb and
normalized to R0F 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
instituted by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 216 a 177 c 183 g 164 t
ORIGIN
Alignment Scores:
Pred. No.: 5.3e-94 Length: 740
Score: 929.50 Matches: 184
Percent Similarity: 95.96% Conservative: 6
Best Local Similarity: 92.93% Mismatches: 8
Query Match: 89.12% Indels: 2
Db: 13 Gaps: 0

```

US-09-782-953-7 (1-198) x BI463566 (1-740)

```

Qy 1 MetAspPheArgAspPheSerTyrAsnPheSerSerLeuIleAlaCysValAlaAsnAsp 20
Db 79 ATGCATTTTAGAAACTTTTAACACAGTTTAGCTCCCTGATTGCGCTGTGGCAACAGT 138
Qy 21 AspValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40
Db 139 GATATCTTCAGCGAAGTGAACACCGGCGCAATTTAGTCCCTCTTTAGGACGTATGAC 198
Qy 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60
Db 199 AAGGACATCACCTTTTCAGTATTTTAAGAGCTTCAAAACGAGTCAGATAAATTCAGCAAC 258
Qy 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
Db 259 CCCTTCTCCGACGACAGATCCAGGCTCCAGCTGCATAAGACTGAGTTTCTGGGAAAGGAA 318
Qy 81 MetLysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProAsn 100
Db 319 ATGAAGTTTATATTTTGTCTCAGACCTTACACATAGGAGCTCACACTGGCTCGGCAAAAT 378
Qy 101 ProAspLysGlnPheLeuIleSerProProAlaSerProProValGlyTyrLysGlnVal 120
Db 379 CCAGACAAGCAGTTTCTGATCTCCCTCTCCGCTCTCCGCGAGTGGATGGAACAAGTG 438
Qy 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro 140
Db 439 GAAGATGCGACCCAGTCATAAACTATGATCTTTATATGCCATCTCCAAGCTGGGGCA 498
Qy 141 GlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160
Db 499 GGGAA-AAGTATGATTTGACGACGACGAGTGCACCACTCCAGCGTGGTGGTCCATGTA 557
Qy 161 CysGluSerAspGlnGluAsnGluGluGluGluMetGluArgMetLysArgPro 180
Db 558 TGTGAGTGTGATCAAGAGAAAGGAGGAGGAGGAAC--ATGGAAAGAAATGAGGAGACCT 615
Qy 181 LysProLysIlelleGlnThrArgArgProGluTyrThrProIleHisLeuSer 198
Db 616 AAGCCAAAATTTATCCAGACGAGGAGCGCGAGTACACCGCATCCACTCAGC 669

```

Search completed: December 14, 2002, 22:28:11
Job time.: 1121.55 secs

THIS PAGE BLANK (USPTO)

A:Cross-references: EMBL:AL355932, GSPDB:GN00116, NCSP:B5022.230
A:Experimental source: BAC clone B5022; strain OK74A

C:Genetics:
A:Gene: NCSP:B5022.230
A:Map position: 6

Query Match 13.5%; Score 140.5; DB 2; Length 315;
Best Local Similarity 34.5%; Pred. No. 8.3e-057; Indels 19; Gaps 6;
Matches 48; Conservative 19; Mismatches 53;

9 NESSLIACANDVSESESTRAKFESLFTYDKTTFQY--FKSFKRVINFSNPLSAD 66
Db NNSL-----DVSADNLQTRIRLISQT---APFAMSPKSPRRIIVTFDEQAIA 118
QY 67 ARLRLHTEFLGKEMKLYFAQTIAIGSS---HLAPPNDKQFLSPSPSPVGVKQ-VE 121
Db 119 VASVMDGEALIGRCGVYFGQPTPIDVSADKHALPDAGKLFISPPSPPHDQRM 178
QY 122 DA--TPVINYDLVLAISKL 138
Db 179 DAPNTVVAHEDLAELAKL 197

RESULT 3
Probable retroelement pol polypeptide [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: D84485
R:Lin. X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.; Neus, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: D84485
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1274 <STO>
A:Cross-references: GB:AE002093; NID:g4309763; PIDN:AAD15532.1; GSPDB:GN00139
C:Genetics:
A:Gene: AC2907400
A:Map position: 2

Query Match 9.4%; Score 98.5; DB 2; Length 1274;
Best Local Similarity 24.8%; Pred. No. 2.3; Indels 82; Gaps 9;
Matches 53; Conservative 28; Mismatches 51;

25 ESETRAKFESLFTYDKTTFQYKSPKRVINFSNPLSADAR-LRLHTEFLGKE--M 81
Db 674 ESRVEAFEGFRGSIENDV-----KQIKELKALADSKSSSYRDMYLAKTOTQODNP 728
QY 82 KLYVAQTIAI---GSSHLAPPND-----KQFLISPA---SPPVGMQ 119
Db 729 KVQTCQTPDVAVKKTNNQFATPSPPSKQADVKGKKTLELLKPKGRGRKSPQPK 788
QY 120 VEDATPVINYDLVLAISKLGGEKYLHA---ATDPTSPVVAVHVCESDQNEEEEM 174
Db 789 VSPMPKTI-----TIKLLKSGQAKKEAEDSDVDVTDVVEYVEMLPESDEDEET 842
QY 175 ERMKRPKPKIIQTR-----PEYTPHLS 198
Db 843 ERLKSVKEIRLKYVLADGRSLINPEAMPHTS 876

RESULT 4
62444
conserved hypothetical protein SPAC136.15c - fission yeast (Schizosaccharomyces pombe)
N:Alternate names: hypothetical protein SPAC24B11.04c
C:Species: Schizosaccharomyces pombe
C:Date: 16-May-1996 #sequence_revision 13-Mar-1997 #text_change 10-Dec-1999
C:Accession: S62444; S62549; T37650; T38331
R:Odell, C.; Bowman, S.

submitted to the EMBL Data Library, October 1995
A:Reference number: S62430
A:Accession: S62444
A:Molecule type: DNA
A:Residues: 1-163 <OD>
A:Cross-references: EMBL:Z54308; NID:g1008985; PIDN:CA91108.1; PID:g1009000
R:Odell, C.; Churcher, C.M.
A:Submitted to the EMBL Data Library, November 1995
A:Reference number: S62546
A:Accession: S62549
A:Molecule type: DNA
A:Residues: 1-163 <OD>
A:Cross-references: EMBL:Z67757; NID:g1061288; PIDN:CA91769.1; PID:g1061292
R:Odell, C.; Bowman, S.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.; Wood, V.

submitted to the EMBL Data Library, October 1995
A:Reference number: Z21734
A:Accession: T37650
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-163 <OD>
A:Cross-references: EMBL:Z54308; PIDN:CA91108.1; GSPDB:GN00066; SPDB:SPAC136.15c
R:Odell, C.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, November 1995
A:Reference number: Z21786
A:Accession: T38331
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-163 <OD>
A:Cross-references: EMBL:Z67757; PIDN:CA91769.1; GSPDB:GN00066; SPDB:SPAC24B11.04c
A:Experimental source: strain 972h-; cosmid c24B11
C:Genetics:
A:Gene: SPAC24B11.04c; SPAC136.15c; SPAC24B11.04c
A:Map position: 1L
C:Superfamily: fission yeast hypothetical protein SPAC136.15c

Query Match 9.2%; Score 95.5; DB 2; Length 163;
Best Local Similarity 32.1%; Pred. No. 0.33; Indels 11; Gaps 4;
Matches 25; Conservative 17; Mismatches 25;

96 LAPNPDKQFLISPPSPVGMQVEDATP--VINYLVAISKLGGEKYLHAATDP 152
Db 85 LQVPKFKGNWLISPPSPVGMQVEDATP--VINYLVAISKLGGEKYLHAATDP 140
QY 153 TPVVAVHVCESDQNEEB 170
Db 141 GPQIVY---SEHNTKE 154

RESULT 5
C84651
hypothetical protein At2g25670 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: C84651
R:Lin. X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.; Neus, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: C84651
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-318 <STO>
A:Cross-references: GB:AE002093; NID:g4874305; PIDN:AAD31367.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g25670
A:Map position: 2

Query Match 9.0%; Score 94; DB 2; Length 318;
Best Local Similarity 25.2%; Pred. No. 1; Indels 30; Gaps 5;
Matches 38; Conservative 19; Mismatches 64;

QY 54 VRINFSNPLSADARLRLHKLHTEFLGKEMKLYFAQTLHGSSHLAPP--NPDKQFLISPPA 111
 Db 13 ITIQSTNLFAALDRKKKKKSKAGSK-----GSSKSRPEPEPEPVVWATTP 62
 QY 112 SPVGVKQVEDATPVINYDLLVAISKLGPEKVELHAATDPTFSPVHVHVCSDQF----- 166
 Db 63 LKVKSWADIDDDDDDDYYATTAPQSGWSTSLPSH-----TDSKDVHVHVESESEDILD 117
 QY 167 -----NNEEEEMERMKRPKPIIQTTRPE 191
 Db 118 EGGDDVEBEQOETEVOVHPEPEV--KKAPE 146

RESULT 6
 T46337
 hypothetical protein DKFZp43402413.1 - human (fragment)
 C:Species: Homo sapiens (man)
 C>Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
 C:Accession: T46337
 R:Koehler, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
 submitted to the Protein Sequence Database, January 2000
 A:Reference number: Z23037
 A:Status: preliminary
 A:Accession: T46337
 A:Molecule type: mRNA
 A:Residues: 1-992 <AAA>
 A:Cross-references: EMBL:AL1137265
 A:Experimental source: adult testis; clone DKFZp43402413
 C:Genetics:
 A:Note: DKFZp43402413.1

Query Match 8.8%; Score 92; DB 2; Length 992;
 Best Local Similarity 22.7%; Pred. No. 6.4;
 Matches 39; Conservative 18; Mismatches 39; Indels 76; Gaps 7;
 QY 96 LAPPNPQKFLISPPAS-----PPVG-----WKQVEDATPVINYDLLVAIS----- 136
 Db 245 LSPPLPHEERAQSPRSLSATEEPPEQPEQPEWKEAE-----LGEDSAASLSQLSLQR 300
 QY 137 -----XKQFG-----EKYELHAATDPTPSV----- 156
 Db 301 EQAPSPPAACEKGKHOHSQAEELGFGQGEAEADPEKVAVSPTTPVSPVSRSTPVPAPPEQ 360
 QY 157 -----VVHVCSQDQNEEEEMERMKRPKPIIQTTRPEYPIH 196
 Db 361 LSEAAKXAMEEVAQVLEQDQRHLLSKQ-EKMQQLREKLCQEBEEELRLH 411

RESULT 7
 T10627
 hypothetical protein T13K14.10 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 15-Oct-1999
 C:Accession: T10627
 R:Bevan, M.; Pohl, T.; Weizensegger, T.; Bancroft, I.; Mayer, K.F.X.; Lemcke
 submitted to the Protein Sequence Database, June 1999
 A:Reference number: Z16991
 A:Accession: T10627
 A:Molecule type: DNA
 A:Residues: 1-1396 <BEV>
 A:Cross-references: EMBL:AL080282; GSPDB:GN00062; ATSP:T13K14.10
 A:Experimental source: cultivar Columbia; BAC clone T13K14
 C:Genetics:
 A:Gene: ATSP:T13K14.10
 A:Map position: 4
 A:Introns: 145/1; 171/2; 196/1; 242/3; 269/3; 298/3; 347/3; 374/3; 414/3; 442/3; 505/2;
 131/3; 1139/3; 1163/3; 1194/3; 1226/3; 1259/3; 1290/3; 1344/3

Query Match 8.3%; Score 87; DB 2; Length 1396;
 Best Local Similarity 22.5%; Pred. No. 27;
 Matches 49; Conservative 24; Mismatches 81; Indels 64; Gaps 7;

QY 18 ANDDFVSESETRAKFESLPRTYDKDTTFQYFKSKVRINFSNPLSADARLRLH----- 72
 Db 1010 AMGDVYFESSKLPKGEYKLYLRHENVELLEKQLQTVFIER--NMGEIRLNLHSEPDG 1067
 QY 73 -----KTEFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPPVGMKQVEDAT 124
 Db 1068 PFTGNGAFKSVLMPGVKEAF-----YLGPPTKDKLPKNTPOGSMVLG----- 1110
 QY 125 PVINYDLLVAISKLGPEKVELH-----AATDPTPSVV----- 158
 Db 1111 -EISYGGKLSFDEKGNPKNDPHRLVKLDAPEDDKKAASAPTCKSVSRLEQEVDRDKI 1169
 QY 159 -HVCSDQNEEEEMERMKRPKPIIOTRPEYTP 195
 Db 1170 KFLGNLQKQETEERSEWRKL-----CTCLKSEYDPDTP 1203

RESULT 8
 S30198
 Na+/H+-exchanging protein - Chinese hamster
 N:Alternate names: Na+/H+ antiporter; Na+/H+ exchanger
 C:Species: Cricetulus griseus (Chinese hamster)
 C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 05-Nov-1999
 C:Accession: S30198; S29541
 R:Counillon, L.; Pouyssegur, J.
 Biochim. Biophys. Acta 1172, 343-345, 1993
 A:Title: Nucleotide sequence of the Chinese hamster Na(+)/H(+) exchanger NHE1.
 A:Reference number: S30198; MUID:93192332; PMID:8383540
 A:Accession: S30198
 A:Molecule type: mRNA
 A:Residues: 1-822 <COU>
 A:Cross-references: EMBL:X68970; NID:G49472; PIDN:CAA48771.1; PID:G49473
 C:Genetics:
 A:Gene: NHE1
 C:Keywords: glycoprotein; ion transport; membrane protein
 F:76,374,414/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 8.3%; Score 86.5; DB 2; Length 822;
 Best Local Similarity 23.7%; Pred. No. 15;
 Matches 40; Conservative 26; Mismatches 64; Indels 39; Gaps 8;
 QY 44 TQYFKSKVRINFSNPLSAA--DARLRLHTEFLGKEMKLYFAQTLH----- 90
 Db 646 TRQRLRSYNRHTL-VADPYEEAWNQMLLRQKARQLEQKMSNYLTVPAAKLDSPTMSRAR 704
 QY 91 IGSSHLA-PPNPDKQFLISPPASPPVGMKQVEDATPVINYDLLVAISKLGPEKVELHAA 149
 Db 705 IQSDPLAYEPKADLPVITIDPASP-----QSPESVDLVNEELKAKV-----LGVN 749

QY 150 TDPTSPVVHVCSQDQNEEEEMERMKRPKPII-----IOTRPEYTP 194
 Db 750 RDPT-----RLTRGDEDEDEDEGVIMWRKEPSSPGTDTVPAPWYSP 793

RESULT 9
 B86414
 hypothetical protein F28N24.14 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001
 C:Accession: B86414
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: B86414; MUID:21016719; PMID:11130712
 A:Accession: B86414
 A:Status: preliminary

Query Match 8.1%; Score 84.5; DB 2; Length 452;
Best Local Similarity 26.3%; Pred. No. 11;
Matches 45; Conservative 19; Mismatches 66; Indels 41; Gaps 8;

QY 2 DPDFSYN-----NFSSLIACVANDVFESESETRAKFESLFTYDKDTTFQY 47
DB 80 DCDASYITPLGTTGGDLDEGNLSFLLTKGNSLFIADAGTVQGVRL-----TFKY 135
QY 48 FSKFKVRINFNSPLSAADARLRLHKTFLGKEMKLYFAQTLHGSSHLAPPNDKQFLI 107
DB 136 FNTL-----FNITPSMAVLPEQR--TSWFLKGVMSVF-----IGHSL--DRVGGGLIV 182
QY 108 SPASPPVGVKQVEDATPVINYDLLVAISKLGPEKVELHAADTPSPVVV 158
DB 183 SPEDYLAKNWIDVQ--PPINNGIMGLIRKLG-----FKPTDFTSSSIL 223

RESULT 13
S51739
transcription repressor ABEP1 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 07-May-1995 #sequence_revision 01-Sep-1995 #text_change 08-Oct-1999
C:Accession: S60227; S51739
R:He, G.P.; Muise, A.; Li, A.W.; Ro, H.S.
Nature 378, 92-96, 1995
A:Title: A eukaryotic transcriptional repressor with carboxypeptidase activity.
A:Reference number: S60227; MUID:96061010; PMID:7477299
A:Accession: S60227
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-719 <HE2>
A:Cross-references: EMBL:X80478; NID:G607131; PIDN:CAA56648.1; PID:G607132

Query Match 8.1%; Score 84.5; DB 2; Length 719;
Best Local Similarity 21.8%; Pred. No. 19;
Matches 31; Conservative 27; Mismatches 37; Indels 47; Gaps 7;

QY 54 VRINFNSPLSAADARLRLHKTFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISP 110
DB 585 LRVDPSRPMTQQORRQORLRMRQML--RRLNSTAGPATSPTP-----ALMPP 637
QY 111 ASP-----PVGWK--QVEDATPVINYDLLVAISKLGPEKVELHAADTP 153
DB 638 PSPTAITLRPWEVLPTTTAGWEESETETYTEVVT-----EFETEYGTGD-- 681

QY 154 PSVVHVHVCSDQNEEEEME 175
DB 682 ----LEVERIEEEEEEMD 699

RESULT 14
A70453
glutamate ammonia ligase adenylyl-transferase - Aquifex aeolicus
C:Species: Aquifex aeolicus
C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999
C:Accession: A70453
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O'V.
Nature 392, 353-358, 1998
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300; MUID:98196666; PMID:9537320
A:Accession: A70453
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-797 <AQF>
A:Cross-references: GB:AE000755; NID:G2984063; PIDN:AAC07604.1; PID:G2984064; GB:AE000695
A:Experimental source: strain VF5
C:Genetics:
A:Gene: glnE

Query Match 8.0%; Score 83.5; DB 2; Length 797;
Best Local Similarity 26.6%; Pred. No. 27;

Matches 51; Conservative 25; Mismatches 63; Indels 53; Gaps 10;

QY 3 FRDFSYN--FSSLIACVANDVFESESETR----AKFESLFTYDK-----DTTFQY 47
DB 427 FTVFSVDYLHLVA--RNPDLVEDALTYRDPFKFEDPKREFEKYRETNLNLSPENLYRR 484
QY 48 FSKFKVRINFNSPLSAADARLRLHKTFLGKEMKLYFAQTLHGSSHLAPPNDKQFLI 107
DB 485 FKTWVEVRIGLVLLKEEDRYEKLKK-----LYISLSTLA-----DFIL 523
QY 108 SPASPPVGVKQVEDATPVINYDLLVAISKLGPEKVELHAADTPSPVVHVHVCSDQEN 167
DB 524 SNL-----WK-----DITLENTALLYALGKLG---SRELNFNSDLDD---LVFAVKSLEEK 568
QY 168 EEEEEEMERMKR 179
DB 569 EKVHEKAKELVR 580

RESULT 15
S37998
probable serine/threonine-specific protein kinase (EC 2.7.1.1) YKL168c - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein YKL632
C:Species: Saccharomyces cerevisiae
C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 24-Sep-1999
C:Accession: S37998; S44593; S38413
R:Vandenbol, M.; Bolle, P.A.; Dion, C.; Portetelle, D.; Hilger, F.
submitted to the Protein Sequence Database, March 1994
A:Reference number: S37976
A:Accession: S37998
A:Molecule type: DNA
A:Residues: 1-734 <VAN>
A:Cross-references: EMBL:Z28168; NID:G486294; PIDN:CAA82010.1; PID:G486295; MIPS:YKL168c
A:Experimental source: strain S288C
R:Vandenbol, M.; Bolle, P.A.; Dion, C.; Portetelle, D.; Hilger, F.
Yeast 10, 25-33, 1994
A:Title: Sequencing and analysis of a 20.5 kb DNA segment located on the left arm of yeast chromosome V.
A:Reference number: S44593
A:Accession: S44593
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-734 <VA2>
A:Cross-references: EMBL:Z26878; NID:G407503; PIDN:CAA81519.1; PID:G407514
A:Experimental source: strain S288C
C:Genetics:
A:Gene: SGD:KKO8
A:Cross-references: SGD:S0001651; MIPS:YKL168c
A:Map position: 11L
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homology
C:Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase
F:420-720/Domain: protein kinase homology <KIN>
F:428-436/Region: protein kinase ATP-binding motif
F:573/Active site: Asp #status predicted

Query Match 7.9%; Score 82.5; DB 2; Length 734;
Best Local Similarity 21.4%; Pred. No. 30;
Matches 42; Conservative 29; Mismatches 64; Indels 61; Gaps 8;

QY 11 SSLIACVANDVFESESETRAKFESLFTYDKDTT-----FOYFKSFKR 53
DB 155 SSVSSCDSSNGTTSSSDSQWAMDDLSDSDNDLTPYRGSKDKILKSKDRAPYNYDDYDK 214
QY 54 VRI---NFSNPLSAADARLRLHKTFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISP 109
DB 215 KALRRATSPNPLPS-----KQFNERLYTRSH-----PDEESLESL 252
QY 110 P--ASPPV-----GWKQVEDATPVINYDLLVAISKLGPEKVELH---AATDTPSVV 157
DB 253 PRFAGADVQCIIIEQNGFKVVEDGSGHEHNKLSGVIAKLEKGNLPAHQSGSLRPRLGIT 312
QY 158 V-----HVCSQDEN 167
DB 313 LSGLFKHKHNECDIEN 328

Sun Dec 15 08:38:43 2002

us-09-782-953-7.rpx

Search completed: December 11, 2002, 11:38:33
Job time : 15.0219 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 11:37:42 ; Search time 24.7083 Seconds
(without alignments)
1651.161 Million cell updates/sec

Title: US-09-782-953-7
Perfect score: 1043
Sequence: 1 MDRPFSYNFSSLIACVAND.....RPKLIQTRRPYTPHLS 198

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_virus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1028	98.6	198	11 Q91WQ4	Q91WQ4 mus musculus
2	839.5	80.5	171	4 Q96R03	Q96R03 homo sapien
3	634.5	60.8	243	11 Q8VIP4	Q8VIP4 mus musculus
4	627.5	60.2	192	11 Q8VIP5	Q8VIP5 mus musculus
5	140.5	13.5	315	3 Q9P5S9	Q9P5S9 neurospora
6	107.5	10.3	249	3 Q9F4A1	Q9F4A1 cryptococcus
7	98.5	9.4	1274	10 Q9ZQK2	Q9ZQK2 arabidopsis
8	94	9.0	318	10 Q9SL96	Q9SL96 arabidopsis
9	94	9.0	318	10 Q9VZ43	Q9VZ43 arabidopsis
10	92.5	8.9	680	17 Q8ZY88	Q8ZY88 pyrobaculum
11	92	8.8	992	4 Q9NTH6	Q9NTH6 homo sapien
12	92	8.8	1455	4 Q9DFV0	Q9DFV0 homo sapien
13	90.5	8.7	4025	4 Q9NR13	Q9NR13 homo sapien
14	88.5	8.5	328	10 Q9LS03	Q9LS03 arabidopsis
15	88	8.4	828	4 Q9BQ15	Q9BQ15 homo sapien
16	87.5	8.4	327	5 Q9XZRI	Q9XZRI clona intes

17	87.5	8.4	1477	5	Q9VB52	Q9VB52 drosophila
18	87	8.3	267	11	Q99JH6	Q99JH6 mus musculus
19	87	8.3	1396	10	Q9SUC7	Q9SUC7 arabidopsis
20	86.5	8.3	1020	10	Q9LP46	Q9LP46 arabidopsis
21	85	8.1	495	10	Q9XIQ8	Q9XIQ8 arabidopsis
22	85	8.1	514	10	Q94C73	Q94C73 arabidopsis
23	85	8.1	560	6	O46411	O46411 bos taurus
24	85	8.1	560	11	Q9D8G6	Q9D8G6 mus musculus
25	84.5	8.1	719	11	O61281	O61281 mus musculus
26	84	8.1	410	5	Q9V3W3	Q9V3W3 drosophila
27	83.5	8.0	797	16	O67651	O67651 aquifex aeo
28	82.5	7.9	611	5	Q8WTE9	Q8WTE9 drosophila
29	82.5	7.9	748	16	Q8YXZ5	Q8YXZ5 anabaena sp
30	82.5	7.9	2015	5	Q9U5Y1	Q9U5Y1 dictyosteli
31	82	7.9	649	16	Q9PJV2	Q9PJV2 chlamydia m
32	82	7.9	697	10	Q9CAI9	Q9CAI9 arabidopsis
33	82	7.9	1128	11	Q88442	Q88442 mus musculus
34	82	7.9	1477	5	O76931	O76931 drosophila
35	81.5	7.8	233	16	Q92780	Q92780 chlamydia p
36	81.5	7.8	528	17	O29668	O29668 archaeoglob
37	81.5	7.8	802	5	O19764	O19764 caenorhabdi
38	81	7.8	1099	4	Q9Y6N4	Q9Y6N4 homo sapien
39	81	7.8	1173	11	O63624	O63624 rattus norv
40	81	7.8	1312	4	Q9NR59	Q9NR59 homo sapien
41	81	7.8	1343	4	Q9H7N4	Q9H7N4 homo sapien
42	80.5	7.7	476	5	O46122	O46122 drosophila
43	80.5	7.7	954	11	O62141	O62141 mus musculus
44	80	7.7	455	10	Q9C752	Q9C752 arabidopsis
45	80	7.7	476	5	Q9VHV8	Q9VHV8 drosophila

ALIGNMENTS

RESULT 1

Q91WQ4 PRELIMINARY; PRT; 198 AA.
ID Q91WQ4;
AC Q91WQ4;
DC 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Down syndrome critical region homolog 1 (human).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RA Strausberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBAJ databases.
DR EMBL; BC013551; AAH13551.1;
SQ SEQUENCE 198 AA; 22851 MW; F017C68F18ACC187 CRC64;

Query Match 98.6%; Score 1028; DB 11; Length 198;
Best Local Similarity 99.0%; Pred. No. 6.9e-92;
Matches 196; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy	1	MDRPFYSNFSSLIACVANDVDFSESETRAKFESLFRITYDKDQTTQYFKSPKRVIRNFNSN	60
Db	1	MHRDFYSNFSSLIACVANDVDFSESETRAKFESLFRITYDKDQTTQYFKSPKRVIRNFNSN	60
Qy	61	PLSAADARLRHKTEFLGKMKLYFAOTLHGSSHLAPNPNKQFLISPPASPPVGWQKV	120
Db	61	PLSAADARLRHKTEFLGKMKLYFAOTLHGSSHLAPNPNKQFLISPPASPPVGWQKV	120
Qy	121	EDATPVINYDLLYAIKSLGPGKEKVELHAATDPTFSSVVHVHVCESDQENEEEMERMKRP	180
Db	121	EDATPVINYDLLYAIKSLGPGKEKVELHAATDPTFSSVVHVHVCESDQENEEEMERMKRP	180
Qy	181	KPKLIQTRRPYTPHLS	198
Db	181	KPKLIQTRRPYTPHLS	198

RESULT 2

Q96R03 PRELIMINARY; PRT; 171 AA.
 ID Q96R03
 AC Q96R03
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Down syndrome critical region protein 1.
 GN DSCR1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 CX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hua F., Wu J., Zhou Y., Zhang B., Peng X., Qiang B., Yuan J.,
 RA Qiang B.,
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF400429; AAK92478.1;
 SQ SEQUENCE 171 AA; 19823 MW; 133001AAEDF0BD9 CRC64;

Query Match 80.5%; Score 839.5; DB 4; Length 171;
 Best Local Similarity 95.3%; Pred. No. 1.1e-73;
 Matches 161; Conservative 3; Mismatches 4; Indels 1; Gaps 1;

QY 30 AKFESLFRYDKDTTFQYKSPKRVINFSNPLSADARLRLHKEFLGKEMKLYFAQTL 89
 DB 4 AKFESLFRYDKDTTFQYKSPKRVINFSNPLSADARLRLHKEFLGKEMKLYFAQTL 63
 QY 90 HIGSSHLAPNDKQFLISPPASPPVGMKQVEDATPVINYLIAISKLGPEKYLHAA 149
 DB 64 HIGSSHLAPNDKQFLISPPASPPVGMKQVEDATPVINYLIAISKLGPEKYLHAA 123
 QY 150 TDPSPVVHVHVGESDQENEEBEMERMKRPKXIIQTRPEYTPHLS 198
 DB 124 TDPSPVVHVHVGESDQENEEBEMERMKRPKXIIQTRPEYTPHLS 171

RESULT 3

Q8VIP4 PRELIMINARY; PRT; 243 AA.
 ID Q8VIP4
 AC Q8VIP4
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Calcineurin inhibitory protein ZAKI-4 beta.
 GN ZAKI-4 BETA.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kanou Y., Miyazaki T., Seo H., Murata Y.,
 RT "calcineurin inhibitory protein ZAKI-4."
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB061525; BAB71956.1;
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF00076; ttm; 1.
 SQ SEQUENCE 243 AA; 27332 MW; 2C08C1A810291851 CRC64;

Query Match 60.8%; Score 634.5; DB 11; Length 243;
 Best Local Similarity 63.9%; Pred. No. 1.4e-53;
 Matches 124; Conservative 22; Mismatches 39; Indels 9; Gaps 2;

QY 5 DFGVSSLLIACVANDVFSESETRAKFESLFRYDKDTTFQYKSPKRVINFSNPLSA 64
 DB 51 DFGVSSLLIACVANDVFSESETRAKFESLFRYDKDTTFQYKSPKRVINFSNPLSA 110
 QY 65 ADARLRRLHKEFLGKEMKLYFAQTLHIGSS----HLAPPNDKQFLISPPASPPVGMKQV 120

DB 111 ARARIELHETQFGKKLKYFAQVQTPETDGDKLHLAPPQPAKQFLISPPSPVGMKPI 170

QY 121 EDATPVINYLIAISKLGPEKYLHAAATDPSPVVHVHVGESDQENEEBEMERMKRP 180
 DB 171 SDATPVINYLIAISKLGPEKYLHAAATDPSPVVHVHVGESDQENEEBEMERMKRP 225

QY 181 KPKXIIQTRPEYTP 194
 DB 226 KPKXIIQTRPEYTP 239

RESULT 4

Q8VIP5 PRELIMINARY; PRT; 192 AA.
 ID Q8VIP5
 AC Q8VIP5
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Calcineurin inhibitory protein ZAKI-4.
 GN ZAKI-4.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kanou Y., Miyazaki T., Seo H., Murata Y.,
 RT "calcineurin inhibitory protein ZAKI-4."
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB061524; BAB71955.1;
 SQ SEQUENCE 192 AA; 21540 MW; B2DB9819F4B36A8 CRC64;

Query Match 60.2%; Score 627.5; DB 11; Length 192;
 Best Local Similarity 63.8%; Pred. No. 4.8e-53;
 Matches 120; Conservative 24; Mismatches 35; Indels 9; Gaps 2;

QY 11 SSLIACVANDVFSESETRAKFESLFRYDKDTTFQYKSPKRVINFSNPLSADARL 70
 DB 6 SSLIACVANDVFSESETRAKFESLFRYDKDTTFQYKSPKRVINFSNPLSADARL 65
 QY 71 LKTEFLGKEMKLYFAQTLHIGSS----HLAPPNDKQFLISPPASPPVGMKQVEDATPV 126
 DB 66 LKTEFLGKEMKLYFAQTLHIGSS----HLAPPNDKQFLISPPASPPVGMKQVEDATPV 125
 QY 127 INYDLIAISKLGPEKYLHAAATDPSPVVHVHVGESDQENEEBEMERMKRPKXIIQ 186
 DB 126 LKCYDLIAISKLGPEKYLHAAATDPSPVVHVHVGESDQENEEBEMERMKRPKXIIQ 180
 QY 187 TRRPEYTP 194
 DB 181 TRRPEYTP 188

RESULT 5

Q9PS59 PRELIMINARY; PRT; 315 AA.
 ID Q9PS59
 AC Q9PS59
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE Nebula related protein.
 GN B5022.230.
 OS Neurospora crassa.
 CC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 CC Sordariales; Sordariaceae; Neurospora.
 CX NCBI_TaxID=5141;
 RN [1]

RP SEQUENCE FROM N.A.
 RA Schulte U., Aign V., Hohnsbein J., Brandt P., Fartmann B., Holland R.,
 RA Nyakatura G., Mewes H.W., Mannhaupt G.,
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 DR [2]
 RN SEQUENCE FROM N.A.

RA Cronin L.A., Shen M., Vanaken S.E., Umayam L., Tallon L.J., Gill J.E.,
 RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
 RA Copenhaver G.P., Preuss D., Nieman W.C., White O., Eisen J.A.,
 RA Salzberg S.L., Fraser C.M., Venter J.C.,
 RA "Sequence and analysis of chromosome 2 of the plant Arabidopsis
 RT thaliana";
 RL Nature 402:761-768 (1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Lin X.;
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AC006053; AAD31367.1;
 SQ SEQUENCE 318 AA; 35128 MW; 2F2E2CD269FFD004 CRC64;

Query Match 9.0%; Score 94; DB 10; Length 318;
 Best Local Similarity 25.2%; Pred. No. 0.55;
 Matches 38; Conservative 19; Mismatches 64; Indels 30; Gaps 5;

OY 54 VRINFSNPLSADARLRLHKTFLGKEMKLYFAQTLHGSSHLAP--NPDKQFLISPPA 111
 DB 13 ITIGSTNLFALDTRKKKKKSDKAGSK-----GSSKSREPEKEPEQVYVWAFPT 62
 OY 112 SPVGVKQVEDATPVINYLVAISKLGPEKYLHAATDPSPVVVAVCSDDQ----- 166
 DB 63 LKYSWADIDDEDDDDYATTAPOSGWSTSLPSH-----TDSKDVHVESESEEDILD 117
 OY 167 -----NEEEEMERMKRPKPKIOTRRPE 191
 DB 118 EGGDDVEEEOEETEVOVHPEPEV--KKAPE 146

RESULT 9
 ID Q8V243 PRELIMINARY; PRT; 318 AA.
 AC Q8V243;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Hypothetical 35.2 kDa protein.
 GN A7G25670.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Eukaryota; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eurosidia II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;

OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Banh J., Chang C.H., Chang E., Dale J.M., Goldsmith A.D.,
 RA Lee J.M., Onodera C.S., Quach H.L., Tang C., Tortum M., Wu H.C.,
 RA Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Cheuk R.,
 RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Kim C., Lam B., Lin J., Meyers M.C., Miranda M.,
 RA Narusaka M., Nguyen A., Palm C.J., Sakurai T., Setou M., Seki M.,
 RA Shim P., Southwick A., Shimozaki K., Davis R.W., Ecker J.R.,
 RA Theologis A.;
 RT "Full length cDNA of gene At2g25670 (GI:15225169).";
 RL Submitted (DEC-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AY065267; AAL38743.1;
 KW Hypothetical protein.
 SQ SEQUENCE 318 AA; 35154 MW; 2F2E2CC278FED004 CRC64;

Query Match 9.0%; Score 94; DB 10; Length 318;
 Best Local Similarity 25.2%; Pred. No. 0.55;
 Matches 38; Conservative 19; Mismatches 64; Indels 30; Gaps 5;

OY 54 VRINFSNPLSADARLRLHKTFLGKEMKLYFAQTLHGSSHLAP--NPDKQFLISPPA 111
 DB 13 ITIGSTNLFALDTRKKKKKSDKAGSK-----GSSKSREPEKEPEQVYVWAFPT 62
 OY 112 SPVGVKQVEDATPVINYLVAISKLGPEKYLHAATDPSPVVVAVCSDDQ----- 166
 DB 63 LKYSWADIDDEDDDDYATTAPOSGWSTSLPSH-----TDSKDVHVESESEEDILD 117

OY 167 -----NEEEEMERMKRPKPKIOTRRPE 191
 DB 118 EGGDDVEEEOEETEVOVHPEPEV--KKAPE 146

RESULT 10
 ID Q8ZY88 PRELIMINARY; PRT; 680 AA.
 AC Q8ZY88;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE DNA replication licensing factor (mcm).
 GN PAE0901.
 OS Pyrobaculum aerophilum.
 OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
 OC Thermoproteaceae; Pyrobaculum.
 OX NCBI_TaxID=13773;

OX NCBI_TaxID=13773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
 RX PubMed=11792869;
 RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
 RA Miller U.H.;
 RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
 RT aerophilum";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989 (2002).
 DR EMBL: AE009792; AL63108.1;
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR001208; MCM.
 DR Pfam: PF00493; MCM_1.
 DR ProDom: PD001041; MCM; 1.
 DR SMART: SM00382; AAA; 1.
 DR SMART: SM00350; MCM; 1.
 DR PROSITE: PS50051; MCM_2; 1.
 KW Complete proteome.
 SQ SEQUENCE 680 AA; 76173 MW; 79409981D729D151 CRC64;

Query Match 8.9%; Score 92.5; DB 17; Length 680;
 Best Local Similarity 22.7%; Pred. No. 2.1;
 Matches 57; Conservative 32; Mismatches 89; Indels 73; Gaps 12;

OY 1 MPRPFSYINFSGLACVANDVDFSESETRAFESLFTYDQDTFOYFKSPKRYINFSN 60
 DB 41 VDFHILMFDSKLADLVVERPKLVLEADKVRIVEKDEETAKALKRFFRR--GS 97
 OY 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHGSSHLAP--NPDKQFLISPPA 111
 DB 98 PLSTV---LRKLRSYIGRLIKIGIVTRQTPPKHFLYKALYRCTGCGYEIELQLERH 154
 OY 96 LAPP-----NPDKQFLISPPSPVGVKQY-----EDATP-----VINDDLVAI 135
 DB 155 VEPAPKCPKCGASKSFYVTELSDYIDQKXIVQERPDLPFGOLPSVVEVLLDLD--V 212
 OY 136 SLGSGEKYELHAATDP-----PSVV-----VAVCSDDQNEEEEMERMKRPK 181
 DB 213 DTVKQEDITSLGVVDLTLSLKKRPPIVTSYIGVAV---DTMKNKLVERI--TKDE 267
 OY 182 PKIQ--TRRP 191
 DB 268 QKLIBSRRP 278

RESULT 11
 ID Q9NTH6 PRELIMINARY; PRT; 992 AA.
 AC Q9NTH6;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE Hypothetical 112.6 kDa protein (Fragment).
 GN DKFP43402413.

```
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RC SEQUENCE FROM N.A.
RA Koenr K., Beyer A., Mewes H.W., Gassenhuber J., Wiemann S.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; ALI37265; CAB70664.1; -
KW Hypothetical protein.
FT NON TER
SQ SEQUENCE 992 AA; 112628 MW; 039DFSB1E97BF02F CRC64;

Query Match
Best Local Similarity 8.8%; Score 92; DB 4; Length 992;
Matches 39; Conservative 18; Mismatches 39; Indels 76; Gaps 7;

QY 96 LAPNPDKQFLISPPAS-----PPVG-----WKQVEDATPVINYDLLVAIS----- 136
Db 245 LSPPLPHEERAQSPRLSATEEPPOGPEQPEWKEAE-----LGDSAAASLSLQLSLQR 300
QY 137 -----KLQPG-----EKYELHAATDPTPSV----- 156
Db 301 EQAPSPPAACEKGQHSQAEEELGPGQEEADEPEEKVAVSPTPPVSPVRSSTPVPAPPEQ 360
QY 157 -----VVHVCSQDQNEEEEEEMRMKRPKPKIIQTRRPETPIH 196
Db 361 LSEAALKAMEEAAVQLEQDQRLHLESQ-EKMQQUREKLCQEEBEILRLH 411

RESULT 12
Q9UPV0 PRELIMINARY; PRT; 1455 AA.
AC Q9UPV0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE KIAA1052 protein.
GN KIAA1052.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RC SEQUENCE FROM N.A.
RA Tanaka A., Kotani H., Nomura N., Ohara O.;
RL MEDLINE=99397452; PubMed=10470851;
RT "Prediction of the coding sequences of unidentified human genes. XIV.
RT for large proteins in vitro."
RL DNA Res. 6:197-205(1999).
DR EMBL; AB028975; BAA83004.1; -
DR InterPro; IPR001202; WW_Rsp5_WWP.
DR Pfam; PF00397; WW; 1.
DR SMART; SM00456; WW; 1.
DR PROSITE; PS00202; WW_DOMAIN 2; 1.
SQ SEQUENCE 1455 AA; 163543 MW; 7F48093100C34819 CRC64;

Query Match
Best Local Similarity 8.8%; Score 92; DB 4; Length 1455;
Matches 39; Conservative 18; Mismatches 39; Indels 76; Gaps 7;

QY 96 LAPNPDKQFLISPPAS-----PPVG-----WKQVEDATPVINYDLLVAIS----- 136
Db 474 LSPPLPHEERAQSPRLSATEEPPOGPEQPEWKEAE-----LGDSAAASLSLQLSLQR 529
QY 137 -----KLQPG-----EKYELHAATDPTPSV----- 156
Db 530 EQAPSPPAACEKGQHSQAEEELGPGQEEADEPEEKVAVSPTPPVSPVRSSTPVPAPPEQ 589
```

```
QY 157 -----VVHVCSQDQNEEEEEEMRMKRPKPKIIQTRRPETPIH 196
Db 590 LSEAALKAMEEAAVQLEQDQRLHLESQ-EKMQQUREKLCQEEBEILRLH 640

RESULT 13
Q9NR13 PRELIMINARY; PRT; 4025 AA.
AC Q9NR13;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE ALR-like protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RC SEQUENCE FROM N.A.
RA Chow V.T.K., Tan Y.C.;
RT "ALR-like protein, a novel human cDNA whose product is homologous to
RL the ALR protein."
CC Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 SET DOMAIN.
DR EMBL; AF264750; AAF74766.2; -
DR InterPro; IPR000194; ATPase_a/bcentre.
DR InterPro; IPR000345; CytC_heme_bind.
DR InterPro; IPR002219; DAG_PE_bind.
DR InterPro; IPR003889; Fyrich_C.
DR InterPro; IPR003888; Fyrich_N.
DR InterPro; IPR000910; HMG_12_box.
DR InterPro; IPR003616; PostSET.
DR InterPro; IPR001214; SET.
DR InterPro; IPR001965; Znf_PHD.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF00505; HMG_box; 1.
DR Pfam; PF00628; PHD; 3.
DR Pfam; PF00856; SET; 1.
DR SMART; SM00109; CL; 1.
DR SMART; SM00542; FYRN; 1.
DR SMART; SM00541; FYRN; 1.
DR SMART; SM00398; HMG; 1.
DR SMART; SM00249; PHD; 4.
DR SMART; SM00508; PostSET; 1.
DR SMART; SM00184; RING; 3.
DR SMART; SM00317; SET; 1.
DR PROSITE; PS00152; ATPASE_ALPHA_BETA; UNKNOWN_1.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
DR PROSITE; PS0280; SET; 1.
SQ SEQUENCE 4025 AA; 443532 MW; 41624149C28E4BDE CRC64;

Query Match
Best Local Similarity 8.7%; Score 90.5; DB 4; Length 4025;
Matches 31; Conservative 15; Mismatches 46; Indels 19; Gaps 4;

QY 87 QTLHIGCSHLAPPDKQF--LISPPASPPVG--WKQVEDATPVINYDLLVAISKLGPCE 142
Db 941 QVFSGSSNSRPPSPMDPYAKVGTFRPPVPGHFSRRNSAAPVENCPTLSSVSR----- 995

QY 143 KYELHAATDPTPSVVHVCSQDQNEEEEEEMRMKRPKPKIIQTRRPET 193
Db 996 PLQMNETTANRPSVRDLCSSTTND-----PYAKPPDTPRPVMT 1036

RESULT 14
Q9LSD3 PRELIMINARY; PRT; 328 AA.
AC Q9LSD3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Genomic DNA, chromosome 3, P1 clone: MOJ10.
```

OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC Eucosida II; Brassicales; Brassicaceae; Arabidopsi.
 OK NCBI_TaxID=3702;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=COLUMBIA;
 RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;
 RL Submitted (Apr-1999) to the EMBL/GenBank/DBJ databases.
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN=COLUMBIA;
 RX MEDLINE=20277480; PubMed=10819329;
 RA Nakamura Y.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
 RT features of the regions of 4,504,864 bp covered by sixty P1 and TAC
 RT clones.";
 RL DNA Rep. 7:131-135(2000).
 DR EMBL; AB026649; BAB01084.1; -;
 DR InterPro; IPR001776; AEROLysin.
 DR PROSITE; PS00274; AEROLYSIN; UNKNOWN 1.
 SQ SEQUENCE 328 AA; 37438 MW; A43A5FE1AD0634AA CRC64;
 Query Match 8.5%; Score 88.5; DB 10; Length 328;
 Best Local Similarity 21.8%; Pred. No. 2;
 Matches 46; Conservative 23; Mismatches 71; Indels 71; Gaps 8;
 QY 46 QYFKSKRVRIN-FS-----NPLSADARLRLHKTFLGKEMKLYFAQTLHGSSHL 96
 DB 41 QYFSVDQDIRTNSFSFYGHTDPNP-SKVEQDFRDEDFCG-----FLAIGTLGT 90
 QY 97 APPNPD-----KQFLISPPASPPVGMKQVEDATPVINYDL 131
 DB 91 DEETPFSAWAEEDATGEIKEMAKLIAKKLDQFLKEYPEDTRSKRKVKSINCEPLQDYDL 150
 QY 132 -----LYAISKLGFEKY-ELHAATDPTPSV--VAVCESD 164
 DB 151 FSSIELTGSGNGRVKKKSLTSLFKRQTVQGEPIEKGHSTRDAIKRVFKLHGASSK 210
 QY 165 QENEEEEEEMERMKRPKXIIOTRPEYTPPI 195
 DB 211 TRNDEDDSMKKKDKLKKVQTCRRKRVHPV 241
 RESULT 15
 Q9BQ15 PRELIMINARY; PRT; 828 AA.
 AC Q9BQ15;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
 DE Hypothetical 89.1 kDa protein.
 GN DKFZP761D221.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OK NCBI_TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=AMYGDALA;
 RX MEDLINE=21154917; PubMed=11320166;
 RA Wiemann S., Weill B., Wellenreuther R., Gaassenhuber J., Glassl S.,
 RA Ansoegge W., Boescher M., Bloecker H., Bauersachs S., Blum H.,
 RA Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Strack N.,
 RA Mewes H.W., Ottenwelder B., Obermayer B., Tampe J., Heubner D.,
 RA Wambutt R., Korn B., Klein M., Poustka A.;
 RT "Towards a Catalog of Human Genes and Proteins: Sequencing and
 RT Analysis of 500 Novel Complete Protein Coding Human cDNAs.";
 RL Genome Res. 11:422-435(2001).
 DR EMBL; AL136561; CAB66496.1; -;
 DR InterPro; IPR000087; Collagen.
 DR InterPro; IPR002965; P_rich_extensn.

DR PRINTS; PR01217; RICHEXTENS.
 KW Hypothetical protein.
 SQ SEQUENCE 828 AA; 89136 MW; E03E971EB7D1DC58 CRC64;
 Query Match 8.4%; Score 88; DB 4; Length 828;
 Best Local Similarity 20.7%; Pred. No. 7.5;
 Matches 50; Conservative 21; Mismatches 79; Indels 92; Gaps 7;
 QY 25 ESETRAFESLFRDYDKDTFOYKSFKRVIRFNSNPLSADARLRLHKTFLGKEMKLY 84
 DB 275 QSATEVEIKELPSINDLSIFGPVLSKVAVN-----AEEKVHV 314
 QY 85 FAQTLHGSSHLAPPNDKQFLISPPASP-----PYG----- 116
 DB 315 FSDT---SPEHVTPELTPREKVVSPATPDNPADSPAPGLGPPGPGPPGNVL 371
 QY 117 -----WKQVEDATPVINYDLLYAISK-----L 138
 DB 372 SPLNLEEVQKVAEQT-FIKDYLETISSPKDGLGORATPPPPPTRYRTVSSPGRS 430
 QY 139 GPGEKYELHAATDPTPSVVAHVCSQDENEEEMERMKRP--KXIIOTRPEYTPPIH 196
 DB 431 GPGPGTSGASSPARPATPLVPCSTTPPPPPPPRPRPKLPKPGVGVSRPSPPIH 490
 QY 197 LS 198
 DB 491 SS 492

Search completed: December 11, 2002, 11:41:06
 Job time : 28.7083 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 11:37:42 ; Search time 30.7184 Seconds
(without alignments)
858.887 Million cell updates/sec

Title: US-09-782-953-7
Perfect score: 1043
Sequence: 1 MDRFRSYNFSLLIACVAND.....RPPKIIQTRRPYTPHLS 198

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database 1:

A_Geneseq_101002:*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1043	100.0	198	23	AAE18912
2	970.5	93.0	197	23	AAE18914
3	920.5	88.3	198	23	AAE18911
4	875	83.9	197	23	AAE18915
5	835.5	80.1	171	20	AAW73898
6	632.5	60.6	197	23	AAE18913
7	623	59.7	234	22	AB31788
8	623	59.7	255	23	AAW50760
9	615.5	59.0	192	23	AAE18916
10	592.5	56.8	241	23	AAE18917

11	591.5	56.7	212	23	AAE18918	Human MCIP associa
12	537	51.5	111	21	ABG01768	Human secreted pro
13	372	35.7	292	22	ABB71467	Drosophila melanog
14	369.5	35.4	142	21	AB58452	Lung cancer associ
15	284	27.2	58	22	AB229495	Peptide #2146 enco
16	284	27.2	58	22	AB234672	Peptide #2178 enco
17	284	27.2	58	22	AB336877	Peptide #4383 enco
18	284	27.2	58	22	AB220081	Protein #2080 enco
19	284	27.2	58	22	AAW55464	Human brain expres
20	284	27.2	58	22	AAW57623	Human brain expres
21	284	27.2	58	22	AAW67851	Human bone marrow
22	284	27.2	58	22	AAW15671	Peptide #2105 enco
23	284	27.2	58	22	AAW17859	Peptide #4293 enco
24	284	27.2	58	22	AAW28174	Peptide #2211 enco
25	284	27.2	58	22	AAW30365	Peptide #4402 enco
26	284	27.2	58	22	AAW03406	Peptide #2088 enco
27	284	27.2	58	22	AAW05506	Peptide #4188 enco
28	284	27.2	58	23	ABG37397	Human peptide enco
29	284	27.2	58	23	ABG39657	Human peptide enco
30	271.5	26.0	56	22	ABW28178	Peptide #859 enco
31	271.5	26.0	56	22	ABW33353	Protein #812 enco
32	271.5	26.0	56	22	ABW18813	Human brain expres
33	271.5	26.0	56	22	AAW54139	Human bone marrow
34	271.5	26.0	56	22	AAW66532	Peptide #839 enco
35	271.5	26.0	56	22	AAW14405	Peptide #855 enco
36	271.5	26.0	56	22	AAW26818	Peptide #814 enco
37	271.5	26.0	56	22	AAW02132	Human peptide enco
38	271.5	26.0	56	23	ABG36184	Human MCIP1 leucin
39	98	9.4	19	23	AAE18921	Arabidopsis thalia
40	94.5	9.1	915	22	ABW12073	Arabidopsis thalia
41	94	9.0	318	21	AAW30396	Arabidopsis thalia
42	94	9.0	318	21	AAW44378	Arabidopsis thalia
43	94	9.0	318	21	AAW49385	Arabidopsis thalia
44	94	9.0	318	21	AAW49386	Arabidopsis thalia
45	92	8.8	954	22	AAU14615	Novel bone marrow

ALIGNMENTS

RESULT 1
AAE18912
ID AAE18912 standard; Protein; 198 AA.
XX
AC AAE18912;
XX
DT 17-MAY-2002 (first entry)
XX
DE Mouse MCIP associated protein #3.
XX
KW Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;
KW heart failure; cardiomyopathy; heart disease; mouse.
XX
OS Mus musculus.
XX
PN WO200204491-A2.
XX
PD 17-JAN-2002.
XX
PF 06-JUL-2001; 2001WO-US21662.
XX
PR 07-JUL-2000; 2000US-216601P.
PR 13-FEB-2001; 2001US-0782953.
XX
PA (TEXA) UNIV TEXAS SYSTEM.
PA (WILL/) WILLIAMS S R.
PA (ROTH/) ROTHERMEL B.
PI Williams SR, Rothermel B;
XX
DR WPI; 2002-179698/23.
DR N-PSDB; AAD30153.
XX

PT	Screening for modulators of muscle calcineurin interacting protein
PT	(MCIP) binding, expression or phosphorylation, useful for treating
PT	cardiac hypertrophy or heart failure, comprises mixing MCIP,
PT	calcineurin and a test compound -
XX	
PS	Diclosure: Page 151-152; 174pp; English.
XX	
CC	The invention relates to muscle calcineurin interacting proteins (MCIPs)
CC	and nucleic acid molecules encoding such proteins. MCIPs form a physical
CC	complex with the catalytic subunit of calcineurin and increased levels
CC	of MCIPs correspond to a reduced ability of calcineurin to stimulate
CC	transcription of certain target genes. The invention also relates to
CC	methods for identifying modulators of MCIP binding, expression or
CC	phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin
CC	may be used for treating cardiac hypertrophy and heart failure.
CC	Antibodies to MCIP can be used in characterizing for determining the
CC	health and diseased tissues and subsequently for determining the
CC	presence or absence of cardiomyopathy or as predictor of heart disease.
CC	The present sequence is mouse MCIP associated protein.
CC	Note: This sequence has been described as human MCIP3 in the
CC	specification, however the sequence seems to be a MCIP associated
CC	protein.
CC	
XX	
SQ	Sequence 198 AA:
Query Match	100.0%; Score 1043; DB 23; Length 198;
Best Local Similarity	100.0%; Pred. No. 7.7e-106; Indels 0; Gaps 0
Matches 198; Conservative	0; Mismatches 0; Indels 0; Gaps 0
OY	1 MDPRDSYNPSSLIACVANDVSESEETAKFESLRTYDKOTTFQYFKSFKVRINFSN 60
DB	1 MDPRDSYNPSSLIACVANDVSESEETAKFESLRTYDKOTTFQYFKSFKVRINFSN 60
OY	61 PLSDADARLRHKTEFLGKEMKLYFAQTLLHIGSSHAPNPDPKQFLISPPASPPVMKKOY 120
DB	61 PLSDADARLRHKTEFLGKEMKLYFAQTLLHIGSSHAPNPDPKQFLISPPASPPVMKKOY 120
OY	121 EDATPVINNDLLAISKLGGEKEYELHAADPTPSVVHVHCSDQENEBEEMEERMKRP 180
DB	121 EDATPVINNDLLAISKLGGEKEYELHAADPTPSVVHVHCSDQENEBEEMEERMKRP 180
OY	181 KPKIOTRREPYTPIHLS 198
DB	181 KPKIOTRREPYTPIHLS 198
RESULT 2	
AAE18914	
ID	AAE18914 standard; Protein; 197 AA.
XX	
AC	AAE18914;
XX	
DT	17-MAY-2002 (first entry)
XX	
DE	Human MCIP associated protein #1.
XX	
KM	Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;
XX	heart failure; cardiomyopathy; heart disease; human.
OS	Homo sapiens.
PN	WO200204491-A2.
XX	
PD	17-JAN-2002.
XX	
PF	06-JUL-2001; 2001MO-US21662.
XX	
PR	07-JUL-2000; 2000US-216601P.
XX	
PA	13-FEB-2001; 2001US-0782953.
XX	
PA	(TEXA) UNIV TEXAS SYSTEM.
XX	
PA	(WILU/) WILLIAMS S R.
XX	
PA	(ROTH/) ROTHELMEL B.

XX	Williams SR, Rothenmel B;
PI	
XX	
DR	MP1: 2002-179698/23.
DR	N-PSDB; AAD30155.
XX	
XX	Screening for modulators of muscle calcineurin interacting protein
PT	(MCIP) binding, expression or phosphorylation, useful for treating
PT	cardiac hypertrophy or heart failure, comprises mixing MCIP,
PT	calcineurin and a test compound -
XX	
PS	Example 1; Page 157-158; 174pp; English.
XX	
CC	The invention relates to muscle calcineurin interacting proteins (MCIPs)
CC	and nucleic acid molecules encoding such proteins. MCIPs form a physical
CC	complex with the catalytic subunit of calcineurin and increased levels
CC	of MCIPs correspond to a reduced ability of calcineurin to stimulate
CC	transcription of certain target genes. The invention also relates to
CC	methods for identifying modulators of MCIP binding, expression or
CC	phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin
CC	may be used for treating cardiac hypertrophy and heart failure.
CC	Antibodies to MCIP can be used in characterising the MCIP content of
CC	healthy and diseased tissues and subsequently for determining the
CC	presence or absence of cardiomyopathy or as predictor of heart disease.
CC	The present sequence is human MCIP associated protein.
CC	Note: This sequence has been described as murine splice variant in
CC	the specification, however the sequence seems to be a MCIP associated
CC	protein.
XX	
SO	Sequence 197 AA;
XX	
Query Match	93.0%; Score 970.5; DB 23; Length 197;
Best Local Similarity	93.4%; Pred. No. 9,3e-100;
Matches 185; Conservative	7; Mismatches 5; Indels 1; Gaps 1
QY	1 MDRFDSYFNSSLIACVANDVFSSEETRAKESLFRIVDKDTTFQYKSPKRVINFSN 60
DB	1 MHRNPNYSFSSLIACVANSIDFSEETRAKESLFRIVDKDTTFQYKSPKRVINFSN 60
QY	61 PLSAARLRIRKHTERLIGKEMKLYPAQTHTIGSSHLAPNPDKPILSPASPVGKQV 120
DB	61 PFSAAARLQIRKHTERLVGKEMKLYPAQTHTIGSSHLAPNPDKQILSPASPVGKQV 120
QY	121 EDATPVINIDLVAISKLGGEKYEELHAATDPTPSVVVHVCSDDENBEEEMERMKP 180
DB	121 EDATPVINIDLVAISKLGGEKYEELHAATDPTPSVVVHVCSDDG-KEEEBEMERMRP 179
QY	181 KPKIIOTRRPEYTPIHLS 198
DB	180 KPKIIOTRRPEYTPIHLS 197
XX	
RESULT 3	
AAE18911	
ID	AAE18911 standard; Protein; 198 AA.
XX	
AC	AAE18911;
XX	
DT	17-MAY-2002 (first entry)
XX	
DE	Human MCIP associated protein #2.
XX	
XX	Muscle calcineurin interacting protein; MCIP, cardiac hypertrophy;
KM	heart failure; cardiomyopathy; heart disease; human.
XX	
OS	Homo sapiens.
XX	
PN	WO200204491-A2.
XX	
PD	17-JAN-2002.
XX	
PF	06-JUL-2001; 2001MO-US21662.
XX	

XX Homo sapiens.
 OS US5869318-A.
 XX
 XX 09-FEB-1999.
 PD
 XX 07-JUN-1996; 96US-0665040.
 PF
 XX 07-JUN-1995; 95ES-0001140.
 PR
 XX (PALU/) PALLEJA X E.
 PA
 XX Puentes JJ, Pallega XE, Pritchard M;
 PI WPI; 1999-152781/13.
 XX N-PSDB; AAX01282.
 DR
 XX
 PT DNA encoding foetal brain proteins - believed to be associated with
 PT Down's syndrome
 PS
 XX Claim 4; Column 15-18; 19pp; English.
 CC This sequence is encoded by the Down's Syndrome critical region 1 (DSCR1)
 CC gene of the invention. The DSCR1 gene was found to be located in the
 CC q22.1-22.2 region of human chromosome 21. An increase in the transient
 CC expression of DSCR1 mRNA in the brains of young rats, compared to
 CC expression levels in the brains of adult rats, suggests an important role
 CC for DSCR1 during the development of the Central Nervous System (CNS), and
 CC that over expression of DSCR1 may be involved in pathogenic abnormalities
 CC of mental retardation and/or heart defects as found in Down's syndrome
 CC patients.
 CC
 SQ Sequence 171 AA;
 Query Match 80.1%; Score 835.5; DB 20; Length 171;
 Best Local Similarity 94.7%; Pred. No. 8.5e-85;
 Matches 160; Conservative 5; Mismatches 3; Indels 1; Gaps 1;
 QY 30 AKESLFRITVDKDTFOYFKSFKRVIRINFSNPISADARLRLKTEFLGEMKLYFAOTL 89
 DB 4 AKESLFRITVDKDTFOYFKSFKRVIRINFSNPISADARLRLKTEFLGEMKLYFAOTL 63
 QY 90 HISSSHLAPNPDKOFLISPPASPVGMKOVEDATPIYINDLYAISKLGPGEYELHAA 149
 DB 64 HISSSHLAPNPDKOFLISPPASPVGMKOVEDATPIYINDLYAISKLGPGEYELHAA 123
 QY 150 TDPTPSVVHVCSDDQNEBEEEMERMRKPKKIITRRPEYTPHLS 198
 DB 124 TDPTPSVVHVCSDDQNEBEEEMERMRKPKKIITRRPEYTPHLS 171
 RESULT 6
 AAB18913
 ID AAB18913 standard; Protein; 197 AA.
 XX
 XX AAB18913;
 AC
 XX 17-MAY-2002 (first entry)
 DT
 XX Mouse MCIP associated protein #4.
 DE
 XX Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;
 KM heart failure; cardiomyopathy; heart disease; mouse.
 KW
 XX Mus musculus.
 OS
 XX MO200204491-A2.
 PN
 XX 17-JAN-2002.
 PD
 XX 06-JUL-2001; 2001WO-US21662.
 PF

PR 07-JUL-2000; 2000US-216601P.
 PR 13-FEB-2001; 2001US-0782953.
 XX
 XX (TEXA) UNIV TEXAS SYSTEM.
 PA (WILLU/) WILLIAMS S R.
 PA (ROTH/) ROTHERMEL B.
 XX
 PI Williams SR, Rothermel B;
 XX
 XX WPI; 2002-179698/23.
 DR N-PSDB; AAD30154.
 DR
 XX Screening for modulators of muscle calcineurin interacting protein
 PT (MCIP) binding, expression or phosphorylation, useful for treating
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,
 PT calcineurin and a test compound -
 PS
 XX Disclosure; Page 154; 174pp; English.
 CC The invention relates to muscle calcineurin interacting proteins (MCIPs)
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical
 CC complex with the catalytic subunit of calcineurin and increased levels
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate
 CC transcription of certain target genes. The invention also relates to
 CC methods for identifying modulators of MCIP binding, expression or
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin
 CC may be used for treating cardiac hypertrophy and heart failure.
 CC Antibodies to MCIP can be used in characterizing the MCIP content of
 CC healthy and diseased tissues and subsequently for determining the
 CC presence or absence of cardiomyopathy or as predictor of heart disease.
 CC The present sequence is mouse MCIP associated protein.
 CC Note: This sequence has been described as mouse MCIP2 encoding DNA in
 CC the specification, however the sequence seems to be a MCIP associated
 CC protein.
 CC
 SQ Sequence 197 AA;
 Query Match 60.6%; Score 632.5; DB 23; Length 197;
 Best Local Similarity 64.4%; Pred. No. 4.5e-62;
 Matches 121; Conservative 24; Mismatches 34; Indels 9; Gaps 2;
 QY 11 SLLICVANDVDVSESEIRAKFBSLFRITVDKDTFOYFKSFKRVIRINFSNPISADARLR 70
 DB 11 SLLICVANDVDVSESEIRAKFBSLFRITVDKDTFOYFKSFKRVIRINFSNPISADARLR 70
 QY 71 LKTEFLGEMKLYFAOTLHIGSS---HLAPNPDKOFLISPPASPVGMKOVEDATPV 126
 DB 71 LKTEFLGEMKLYFAOTLHIGSS---HLAPNPDKOFLISPPASPVGMKOVEDATPV 130
 QY 127 INVDLYAISKLGPGEYELHAAIDPTPSVVHVCSDDQNEBEEEMERMRKPKKIITQ 186
 DB 131 INVDLYAISKLGPGEYELHAAIDPTPSVVHVCSDDQNEBEEEMERMRKPKKIITQ 185
 QY 187 TRRPEYTP 194
 DB 186 TRRPEYTP 193
 RESULT 7
 AAB31788
 ID AAB31788 standard; Protein; 234 AA.
 XX
 XX AAB31788;
 AC
 XX 30-APR-2001 (first entry)
 DT
 XX Amino acid sequence of a human detoxification protein.
 DE
 XX Human; detoxification protein; DCTX; cancer; leukemia; melanoma;
 KM adenocarcinoma; autoimmune disorder; inflammatory disorder;
 KM rheumatoid arthritis; asthma; atherosclerosis; autoimmune thyroiditis;
 KM psoriasis; ulcerative colitis; infection; cell proliferative disorder;
 KM actinic keratosis; arteriosclerosis; cirrhosis; hepatitis.

SQ Sequence 255 AA: 59.7%; Score 623; DB 23; Length 255;
 Query Match Similarity 62.4%; Pred. No. 7,6e-61;
 Best Local Similarity 62.4%; Pred. No. 7,6e-61;
 Matches 123; Conservative 21; Mismatches 41; Indels 12; Gaps 3;

QY 2 DPRDSYNSFSLIACVANDVSESETRAKFESLFRITDKDTTFQYFSPKRVIRINFSNP 61
 DB 63 DPNDDP---NSLIFACVNHQSVFEGESKEKPEGLFRITDDCTFQLFSPFRVRIRINFSNP 119
 QY 62 LSAADARLRLHKTETELGKEMKLYFAQTLLHIGSS---HLAPPNPKQPLISPPASPPVGM 117
 DB 120 KSAARARIELHETQFRGKKLKLKYPAQVQPTPTDGDKLHLAPQAKQFLISPPSSPPVGM 179
 QY 118 KQVEDATPVINDLYLAISKLGEGEYELHAATDPTPSVVVHVCSDDQNEEBEEMERBM 177
 DB 180 QVINDATPVINDLYLAIVAKLGEGEYELHAGTSTPSVVVHVCSDDIEEDPK----- 234

QY 178 KRPKPKIOTRRPEYTP 194
 DB 235 TSPKPKIOTRRPEYTP 251

RESULT 9
 AAE18916
 ID AAE18916 standard; Protein; 192 AA.
 AC AAE18916;
 XX 17-MAY-2002 (first entry)
 DT Human MCIP associated protein #3.
 DE Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;
 XX heart failure; cardiomyopathy; heart disease; human.
 KW Homo sapiens.
 OS WO200204491-A2.
 PN 17-JAN-2002.
 PD 06-JUL-2001; 2001WO-US21662.
 PF 07-JUL-2000; 2000US-216601P.
 PR 13-FEB-2001; 2001US-0782953.
 PA (TEXA) UNIV TEXAS SYSTEM.
 PA (WILL/) WILLIAMS S R.
 PA (ROTH/) ROTHERMEL B.
 PI Williams SR, Rothermel B;
 XX MPI; 2002-179698/23.
 DR N-PSDB; AAD30157.
 XX Screening for modulators of muscle calcineurin interacting protein
 PT (MCIP) binding, expression or phosphorylation, useful for treating
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,
 PT calcineurin and a test compound -
 XX Example 1; Page 165-166; 174pp; English.

CC presence or absence of cardiomyopathy or as predictor of heart disease.
 CC The present sequence is human MCIP associated protein.
 CC Note: This sequence has been described as human MCIP splice variant in
 CC specification, however the sequence seems to be a MCIP associated
 CC protein.

SQ Sequence 192 AA: 59.0%; Score 615.5; DB 23; Length 192;
 Query Match Similarity 63.3%; Pred. No. 3,4e-60;
 Best Local Similarity 63.3%; Pred. No. 3,4e-60;
 Matches 119; Conservative 23; Mismatches 37; Indels 9; Gaps 2;

QY 11 SSLIACVANDVSESETRAKFESLFRITDKDTTFQYFSPKRVIRINFSNPISAADARLR 70
 DB 6 STIACVADVVEFTNOEVEKEKPEGLFRITDDCTFQLFSPFRVRIRINFSNPISAADARLR 65
 QY 71 LKHEFEGKEMKLYFAQTLLHIGSS---HLAPPNPKQPLISPPASPPVGMKQVEDATPV 126
 DB 66 LHETQFRGKKLKLKYPAQVQPTPTDGDKLHLAPQAKQFLISPPSSPPVGMQVINDATPV 125
 QY 127 INVDLYLAISKLGEGEYELHAATDPTPSVVVHVCSDDQNEEBEEMERBMKRPKPKIIQ 186
 DB 126 LNYDLLYAVAKLGEGEYELHAGTSTPSVVVHVCSDDIEEDPK-----TSPKPKIIQ 180
 QY 187 TRRPEYTP 194
 DB 181 TRRPEYTP 188

RESULT 10
 AAE18917
 ID AAE18917 standard; Protein; 241 AA.
 AC AAE18917;
 XX 17-MAY-2002 (first entry)
 DT Human MCIP associated protein #4.
 DE Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;
 XX heart failure; cardiomyopathy; heart disease; human.
 KW Homo sapiens.
 OS WO200204491-A2.
 PN 17-JAN-2002.
 PD 06-JUL-2001; 2001WO-US21662.
 PF 07-JUL-2000; 2000US-216601P.
 PR 13-FEB-2001; 2001US-0782953.
 PA (TEXA) UNIV TEXAS SYSTEM.
 PA (WILL/) WILLIAMS S R.
 PA (ROTH/) ROTHERMEL B.
 PI Williams SR, Rothermel B;
 XX MPI; 2002-179698/23.
 DR N-PSDB; AAD30158.
 XX Screening for modulators of muscle calcineurin interacting protein
 PT (MCIP) binding, expression or phosphorylation, useful for treating
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,
 PT calcineurin and a test compound -
 XX Example 1; Page 168-169; 174pp; English.

CC The invention relates to muscle calcineurin interacting proteins (MCIPs)
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical
 CC complex with the catalytic subunit of calcineurin and increased levels
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate
 CC transcription of certain target genes. The invention also relates to
 CC methods for identifying modulators of MCIP binding, expression or
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin
 CC may be used for treating cardiac hypertrophy and heart failure.
 CC Antibodies to MCIP can be used in characterizing the MCIP content of
 CC healthy and diseased tissues and subsequently for determining the

CC transcription of certain target genes. The invention also relates to
 CC methods for identifying modulators of MCIP binding, expression or
 CC phosphorylation. Inhibitors or promoters of MCIP binding, expression or
 CC may be used for treating cardiac hypertrophy and heart failure.
 CC Antibodies to MCIP can be used in characterizing the MCIP content of
 CC healthy and diseased tissues and subsequently for determining the
 CC presence or absence of cardiomyopathy or as predictor of heart disease.
 CC The present sequence is human MCIP associated protein.
 CC Note: This sequence has been described as splice variant of MCIP1
 CC initiated by exon 4 in the specification, however the sequence seems
 CC to be a MCIP associated protein.
 XX
 SQ Sequence 241 AA;
 Query Match 56.8%; Score 592.5; DB 23; Length 241;
 Best Local Similarity 61.5%; Pred. No. 1.8e-57;
 Matches 120; Conservative 19; Mismatches 43; Indels 13; Gaps 3;
 QY 1 MDRFDSYNFSLIACVANDVSESETRAKFESLRTYDKDTTFQYFKSPKRVINFSN 60
 DB 40 MDLSLP---TSLFACSVHEAVFEAREQKERFEALFTYDDQVTFQLFKSPRVINFSK 96
 QY 61 PLSAADARLRHKTFLGKEMKLYFAQTLHG-----SSHLAPPNDPKQFLISPPASPPVG 116
 DB 97 PEAARARIELHETDFNGOKLKYFAQVMSGSEVRDKSYLLPPQVQKFLISPPASPPVG 156
 QY 117 WKQVEDATPVINYDLLYVAISKLGPGEKVELHAATDPTSPVVHVHVCSDQNEEEEMER 176
 DB 157 WKQBEDAMPVINYDLLCAVSKLGPGEKVELHAGTSTFSPVVHVHVCSETEEEEEE----- 210
 QY 177 MKRPKPIIOTRRPE 191
 DB 211 TKNPKQIAQTRRPD 225
 RESULT 11
 AAE18918
 ID AAE18918 standard; Protein; 212 AA.
 XX
 AC AAE18918;
 XX
 DT 17-MAY-2002 (first entry)
 XX
 DE Human MCIP associated protein #5.
 XX
 KW Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;
 KW heart failure; cardiomyopathy; heart disease; human.
 XX
 OS Homo sapiens.
 XX
 PN W0200204491-A2.
 XX
 PD 17-JAN-2002.
 XX
 PF 06-JUL-2001; 2001WO-US21662.
 XX
 PR 07-JUL-2000; 2000US-216601P.
 PR 13-FEB-2001; 2001US-0782953.
 XX
 PA (TEXA) UNIV TEXAS SYSTEM.
 PA (WILL/) WILLIAMS S R.
 PA (ROTH/) ROTHERMEL B.
 XX
 PI Williams SR, Rothermel B;
 XX
 DR WPI; 2002-179698/23.
 DR N-PSDB; AAD30159.
 XX
 PT Screening for modulators of muscle calcineurin interacting protein
 PT (MCIP) binding, expression or phosphorylation, useful for treating
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,
 PT calcineurin and a test compound -
 XX

PS Disclosure; Page 171-172; 174pp; English.
 XX
 CC The invention relates to muscle calcineurin interacting proteins (MCIPs)
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical
 CC complex with the catalytic subunit of calcineurin and increased levels
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate
 CC transcription of certain target genes. The invention also relates to
 CC methods for identifying modulators of MCIP binding, expression or
 CC phosphorylation. Inhibitors or promoters of MCIP binding, expression or
 CC may be used for treating cardiac hypertrophy and heart failure.
 CC Antibodies to MCIP can be used in characterizing the MCIP content of
 CC healthy and diseased tissues and subsequently for determining the
 CC presence or absence of cardiomyopathy or as predictor of heart disease.
 CC The present sequence is human MCIP associated protein described in the
 CC invention.
 XX
 SQ Sequence 212 AA;
 Query Match 56.7%; Score 591.5; DB 23; Length 212;
 Best Local Similarity 62.3%; Pred. No. 1.9e-57;
 Matches 119; Conservative 19; Mismatches 38; Indels 15; Gaps 3;
 QY 1 MDRFDSYNFSLIACVANDVSESETRAKFESLRTYDKDTTFQYFKSPKRVINFSN 60
 DB 21 MDLSLP---TSLFACSVHEAVFEAREQKERFEALFTYDDQVTFQLFKSPRVINFSK 77
 QY 61 PLSAADARLRHKTFLGKEMKLYFAQTLHGSSHLAPPNDPKQFLISPPASPPVGKQV 120
 DB 78 PFAARARIELHETDFNGOKLKYFAQ-----SYLLPPQVQKFLISPPASPPVGKQK 131
 QY 121 EDATPVINYDLLYVAISKLGPGEKVELHAATDPTSPVVHVHVCSDQNEEEEMERMKRP 180
 DB 132 EDAMPVINYDLLCAVSKLGPGEKVELHAGTSTFSPVVHVHVCSETEEEEEE-----TKNP 185
 QY 181 KPKIOTRRPE 191
 DB 186 KQKIAQTRRPD 196
 RESULT 12
 AAG01768
 ID AAG01768 standard; Protein; 111 AA.
 XX
 AC AAG01768;
 XX
 DT 06-OCT-2000 (first entry)
 XX
 DE Human secreted protein, SEQ ID NO: 5849.
 XX
 KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW Gene therapy; chromosome mapping.
 XX
 OS Homo sapiens.
 XX
 PN EPI033401-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 21-FEB-2000; 2000EP-0200610.
 XX
 PR 26-FEB-1999; 99US-0122487.
 XX
 PA (GEST) GENSET.
 XX
 PI Dumas Milne Edwards J, Duclert A, Giordano J;
 XX
 DR WPI; 2000-500381/45.
 DR N-PSDB; AAC01774.
 XX
 PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 XX

PS Claim 13, SEQ ID 5849; 71pp + CD-ROM; English.

XX The present sequence is a polypeptide encoded by one of a large number
CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
CC were prepared from total human RNAs or polyA+ RNAs derived from 30
CC different tissues. EST sequences usually correspond mainly to the 3'
CC untranslated region (UTR) of the mRNA because they are often obtained
CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
CC those cases where longer cDNA sequences have been obtained, the full 5'
CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
CC DNA and can therefore be used to obtain full length cDNAs and genomic
CC DNA. 5' ESTs are also used in diagnostic, forensic, gene therapy and
CC regulatory mapping procedures. They are used to obtain upstream
CC regulatory sequences and to design expression and secretion vectors.

XX

XX Sequence 111 AA:

Query Match 51.5%; Score 537; DB 21; Length 111;
Best Local Similarity 91.9%; Pred No. 8.7e-52;
Matches 102; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 MPRDPSYFSSLIACVANDVSESETRAKFESLFTYDKDTTFQYFKSKVRINFSN 60
DB 1 MHRNPNYSSSLIACVANDVSESETRAKFESLFTYDKDTTFQYFKSKVRINFSN 60

QY 61 PLSADARLRLKTEFLGKEMKLYFAQTLLHIGSSHLAPNPKOFLISPPA 111
DB 61 PFSADARLRLKTEFLGKEMKLYFAQTLLHIGSSHLAPNPKOFLISPPA 111

Db

RESULT 13

AB571467

XX ABB71467 standard; Protein; 292 AA.

XX

AC ABB71467;

XX

DT 26-MAR-2002 (first entry)

XX

DE Drosophila melanogaster polypeptide SEQ ID NO 41193.

XX

KM Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.

XX

OS Drosophila melanogaster.

XX

PN WO200171042-A2.

XX

PD 27-SEP-2001.

XX

PF 23-MAR-2001; 2001WO-US09231.

XX

PR 23-MAR-2000; 2000US-191637P.

XX

PR 11-JUL-2000; 2000US-0614150.

XX

PA (PEKE) PE CORP NY.

XX

PI Venter JC, Adams M, Li PMD, Myers EW;

XX

DR WPI; 2001-656860/75.

XX

DR N-PSDB; ABL15570.

XX

PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -

XX

PS Disclosure; SEQ ID NO 41193; 21pp + Sequence Listing; English.

XX

CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutic and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (AB57373-AB57072).

CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcc_sequences.

XX

XX Sequence 292 AA:

Query Match 35.7%; Score 372; DB 22; Length 292;
Best Local Similarity 41.0%; Pred No. 9e-33;
Matches 75; Conservative 42; Mismatches 56; Indels 10; Gaps 4;

QY 11 SSLIACVANDVSESETRAKFESLFTYDKDTTFQYFKSKVRINFSNPLSADARL 70
DB 116 TSLIVNHSEVFANPFLKNAMELFRFTFSESATFQWLRFRRLRVYDVAIAANARIK 175

QY 71 LKTEFLGKE-MKLYFAQTLL-HIGSSHLAPNPKOFLISPPAPVGMQVADATPVIN 128
DB 176 LHOYEFNKKQVITCYFAQPTVPVSNKNLQPPAPVKOFLISPPAPGAPREGEPLVN 235

QY 129 YDLIYATSKLQGEKYEELHAATDPTPSVVHVHVCESDQENEEBEMERMKRPXKIOTR 188
DB 236 HDLALALASLTGSHLHPOSBDQPAIIVHTML-----AETGPGLVQVAP---IVQTK 287

QY 189 RPE 191
DB 288 CPE 290

Db

RESULT 14

AAB58452

XX AAB58452 standard; Protein; 142 AA.

XX

AC AAB58452;

XX

DT 14-MAR-2001 (first entry)

XX

DE Lung cancer associated polypeptide sequence SEQ ID 790.

XX

KM Human; lung cancer associated protein; neuroprotective; cyrostatic;
KW cardioactive; immunomodulatory; muscular active; vulnery;
KW gastrointestinal; nephrotoxic; antiinfective; gynecological;
KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
KW proliferative disorder; wound healing; infectious disease.

XX

OS Homo sapiens.

XX

PN WO200055180-A2.

XX

PD 21-SEP-2000.

XX

PF 08-MAR-2000; 2000WO-US05918.

XX

PR 12-MAR-1999; 99US-0124270.

XX

PA (HUMA-) HUMAN GENOME SCT INC.
(ROSE/) ROSEN C A.

XX

PI Ruben SM;

XX

DR WPI; 2000-587514/55.

XX

DR N-PSDB; AAF18328.

XX

PT Lung cancer associated gene sequences, referred to as lung cancer
PT antigens, useful for treatment, prevention, and diagnosis of disorders
PT such as lung cancer -

XX

PS Claim 11; Page 1325-1326; 1425pp; English.

XX

XX Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer
XX associated proteins represented in AAB58106 - AAB58548. Lung cancer
XX associated proteins and polynucleotide sequences, their agonists, and

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 11:37:42 ; Search time 7.3457 Seconds
(without alignments)
1117.976 Million cell updates/sec

Title: US-09-782-953-7

Perfect score: 1043

Sequence: 1 MDRFRSNNFSSLIACVAND.....RPKPKIIQTRRPTPIHLS 198

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	992.5	95.2	197	1 CCPI CRIGR	O35847 cricetulus
2	912.5	87.5	198	1 CCPI MOUSE	Q9jhg6 mus musculus
3	875	83.9	197	1 CCPI HUMAN	P53805 homo sapien
4	632.5	60.6	197	1 CCP2 MOUSE	Q9jhg2 mus musculus
5	615.5	59.0	197	1 CCP2 HUMAN	Q14206 homo sapien
6	594.5	57.0	239	1 CCP3 MOUSE	Q9jkk0 mus musculus
7	592.5	56.8	241	1 CCP3 HUMAN	Q9ukx8 homo sapien
8	372	35.7	232	1 NL4 DROME	Q9x218 drosophila
9	351	33.7	207	1 CCPL CAEL	P53806 caenorhabdi
10	95.5	9.2	163	1 YA9F SCHPO	Q09791 schizosacch
11	86.5	8.3	822	1 NAH1 CRIGR	P48761 cricetulus
12	85.5	8.2	561	1 SNTC HUMAN	P49902 homo sapien
13	82.5	7.9	734	1 KQ08 YEAST	P36004 saccharomyc
14	82	7.9	370	1 CR42 BACSH	P05516 bacillus sp
15	81.5	7.8	1790	1 SEPA EMENI	P78621 emericella
16	81	7.8	150	1 ALI5 HEVER	Q39967 hevea bras
17	81	7.8	2339	1 CCAB HUMAN	Q00975 homo sapien
18	80.5	7.7	360	1 VTFA TTVI	P19274 thermoprote
19	80	7.7	370	1 CR41 BACSH	P06575 bacillus sp
20	79.5	7.6	452	1 CN17 DICDI	P12019 dictyosteli
21	79.5	7.6	559	1 ECM1 MOUSE	Q61508 mus musculus
22	79.5	7.6	587	1 RGP1 HUMAN	P46060 mus sapien
23	78.5	7.5	201	1 MAP2 SCHPO	Q09180 schizosacch
24	78.5	7.5	211	1 RCN1 YEAST	P36054 saccharomyc
25	78	7.5	370	1 CR43 BACSH	P12963 bacillus sp
26	78	7.5	467	1 INVO MOUSE	P48997 mus musculus
27	78	7.5	1355	1 SALM DROME	P39770 drosophila
28	77.5	7.4	592	1 ABP1 YEAST	P15891 saccharomyc
29	77.5	7.4	627	1 DNK1 GUITH	P29215 guillardia
30	77.5	7.4	630	1 YCF2 OENVI	P31569 oenothera v
31	77.5	7.4	721	1 YCF2 OENPI	P31568 oenothera p
32	77	7.4	681	1 DP3X MYCPN	P75177 mycoplasma
33	77	7.4	691	1 CYG3 BOVIN	P19687 bos taurus

ALIGNMENTS

RESULT 1

```

CCPI_CRIGR
ID CCPI_CRIGR STANDARD; PRT; 197 AA.
AC O35847;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Calcipressin 1 (Down syndrome critical region protein 1 homolog)
DE (Oxidative-induced protein Adapt78).
GN DSCR1 OR ADAPT78.
OS Cricetulus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetulus.
OX NCBI_TaxID=10029;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97329095; PubMed=9185608;
RA Crawford D.R., Leahy K.P., Abramova N., Lan L., Wang Y., Davies K.J.;
RT "Hamster adapt78 mRNA is a Down syndrome critical region homologue
RT that is inducible by oxidative stress.";
RL Arch. Biochem. Biophys. 342:6-12(1997).
CC -!- FUNCTION: Inhibits calcineurin-dependent transcriptional responses
CC by binding to the catalytic domain of calcineurin A. Could play a
CC role during central nervous system development (By similarity).
CC -!- INDUCTION: OXIDANT-INDUCIBLE.
CC -!- SIMILARITY: BELONGS TO THE DSCR1 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL; U60263; AB68517.1; -
SQ SEQUENCE 197 AA; 22674 MW; 0BFC68FA55826439 CRC64;

```

```

Query Match          95.2%; Score 992.5; DB 1; Length 197;
Best Local Similarity 96.5%; Pred. No. 3.5e-78;
Matches 191; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 1 MDRFRSNNFSSLIACVANDVDFSESETRAKFESLFRDYKDTTFQYFKSKFRVRINFSN 60
DB 1 MHFRDNNFSSLIACVANGDVFSESETRAKFESLFRDYKDTTFQYFKSKFRVRINFSN 60
QY 61 PLSAADARLRHKTEFLGKEMKLYFAQTLHGSSHLAPPNDPKQFLSPASPVPVGVKQV 120
DB 61 PLSAADARLRHKTEFLGKEMKLYFAQTLHGSSHLAPPNDPKQFLSPASPVPVGVKQV 120
QY 121 EDATPVINYDLLYALSKLGPCKEYELHAATDPTPSVVHVHVCESDQENEEEEEEMMKRP 180
DB 121 EDATPVINYDLLYALSKLGPCKEYELHAATDPTPSVVHVHVCESDQEN-EBEEEEMMKRP 179
QY 181 KPKIIQTRRPTPIHLS 198

```


RA Ectivill X.;

RT "A new human gene from the Down syndrome critical region encodes a

RL proline-rich protein highly expressed in fetal brain and heart.";

RN Hum. Mol. Genet. 4:1935-1944(1995).

RP [2]

RA SEQUENCE FROM N.A.; AND ALTERNATIVE SPLICING.

RP MEDLINE=97468152; PubMed=9325060;

RA Fuentes J.-J., Pritchard M.A., Estivill X.;

RT "Genomic organization, alternative splicing, and expression patterns

RT of the DSCR1 (Down syndrome candidate region 1) gene.";

RL Genomics 44:358-361(1997).

RN [3]

RA SEQUENCE FROM N.A. (ISOFORM 1).

RP TISSUE=Lung;

RC Strausberg R.;

RA Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.

RL [4]

RA SEQUENCE OF 1-164 FROM N.A. (ISOFORM 2).

RP TISSUE=Breast;

RA Crawford D.R., Leahy K.P., Davies K.J.A.;

RT "Adapt78, a calcium and oxidant-inducible RNA.";

RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.

RN [5]

RA CHARACTERIZATION.

RP MEDLINE=20320698; PubMed=10861295;

RA Fuentes J.J., Genesca L., Kingsbury T.J., Cunningham K.W.,

RA Perez-Riba M., Estivill X., de la Luna S.;

RT "DSCR1, overexpressed in Down syndrome, is an inhibitor of

RT calcineurin-mediated signaling pathways.";

RL Hum. Mol. Genet. 9:1681-1690(2000).

CC -1- FUNCTION: Inhibits calcineurin-dependent transcriptional responses

CC by binding to the catalytic domain of calcineurin A. Could play a

CC role during central nervous system development.

CC -1- ALTERNATIVE PRODUCTS: 4 ISOFORMS; 1 (SHOWN HERE), 2, 3 AND 4; ARE

CC PRODUCED BY ALTERNATIVE SPLICING.

CC -1- TISSUE SPECIFICITY: Highly expressed heart, brain and skeletal

CC muscle. Also expressed in all other tissues.

CC -1- INDUCTION: By calcium.

CC -1- SIMILARITY: BELONGS TO THE DSCR1 FAMILY.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC

DR EMBL; U28833; AAB81557.1; -

DR EMBL; U85265; AAB84370.1; -

DR EMBL; U85266; AAB84371.2; -

DR EMBL; U85267; AAB84372.1; -

DR EMBL; BC002864; AAB02864.1; -

DR EMBL; U53821; AAF21218.1; -

DR Genew; HGNC:3040; DSCR1.

DR MIM: 602917; -

KW Alternative splicing.

FT DOMAIN 166 174

FT VARSPPLIC 1 28

FT

FT POLY-GLU.

FT MEVDLQDLPSATIACHLDPVFDGLC -> MIFRNFNYS

FT FSLLIACVANSDFISET (IN ISOFORM 2).

FT MEVDLQDLPSATIACHLDPVFDGLC -> MYV (IN

FT ISOFORM 3).

FT MISSING (IN ISOFORM 4).

FT CONFLICT 159 159 H -> Q (IN REF. 4).

FT SEQUENCE 197 AA; 22637 MW; 1BD426BB88167E6C CRC64;

SQ

Query Match 83.9%; Score 875; DB 1; Length 197;

Best Local Similarity 86.4%; Pred. No. 48-68;

Matches 171; Conservative 8; Mismatches 15; Indels 4; Gaps 2;

QY 1 MDRFDSYFNFSLLIACVANDVSESETRAKFESLRTYDKDTTFQYFKSKVRINFSN 60

DB 4 VDLQDLP---SATIACHLDPVFDGLCRAKFSLSLRTYDKDTTFQYFKSKVRINFSN 60

QY 61 PLSAADARLRLHKTEFLGKEMKLYFAQTLLHIGSSHLAPPNDKQFLISPPAPPGVKQV 120

DB 61 PLSAADARLRLHKTEFLGKEMKLYFAQTLLHIGSSHLAPPNDKQFLISPPAPPGVKQV 120

QY 121 EDATPVINDLLYAIKLGEGEKYELHAATDPTPSVVHVCSQDQNEEEEEEMERMRP 180

DB 121 EDATPVINDLLYAIKLGEGEKYELHAATDPTPSVVHVCSQDQNEEEEEEMERMRP 179

QY 181 KPKIOTRRPEYTPHLS 198

DB 180 KPKIOTRRPEYTPHLS 197

RESULT 4

CCP2_MOUSE

ID CCP2_MOUSE STANDARD; PRT; 197 AA.

AC Q9JHG2;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DE Calcipressin 2 (Down syndrome candidate region 1-like protein 1)

DE (Myocyte-enriched calcineurin interacting protein 2) (MCIP2).

GN DSCR1L1.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=20187590; PubMed=10722714;

RA Rothermel B., Vega R.B., Yang J., Wu H., Bassel-Duby R.,

RA Williams R.S.;

RT "A protein encoded within the Down syndrome critical region is

RT enriched in striated muscles and inhibits calcineurin signaling.";

RL J. Biol. Chem. 275:8719-8725(2000).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=BALB/c; TISSUE=Brain;

RX MEDLINE=20534792; PubMed=11080588;

RA Strippoli P., Petri M., Lenzi L., Carinci P., Zannotti M.;

RT "The murine DSCR1-like (Down syndrome candidate region 1) gene family:

RT conserved synteny with the human orthologous genes.";

RL Gene 257:223-232(2000).

CC -1- FUNCTION: Inhibits calcineurin-dependent transcriptional responses

CC by binding to the catalytic domain of calcineurin A. Could play a

CC role during central nervous system development.

CC -1- TISSUE SPECIFICITY: Highest expression in heart, skeletal muscle

CC and brain. Lower expression in all other tissues.

CC -1- SIMILARITY: BELONGS TO THE DSCR1 FAMILY.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC

DR EMBL; AF237791; AAF63487.1; -

DR EMBL; AF237887; AAF62538.1; -

DR MGD; MGI:1858219; Dscr1l1.

SQ SEQUENCE 197 AA; 22025 MW; CB1306B7B03E70F1 CRC64;

Query Match 60.6%; Score 632.5; DB 1; Length 197;

Best Local Similarity 64.4%; Pred. No. 2.3e-47;

Matches 121; Conservative 24; Mismatches 34; Indels 9; Gaps 2;

QY 11 SSLIACVANDVSESETRAKFESLRTYDKDTTFQYFKSKVRINFSNPLSADARLR 70

DB 11 STLVCVVDVEVFTNQEKVEKFEGLFRYDCEVTFQLFSPRRVIRNFSHPKSAARARIE 70

QY 71 LHKTEFLGKEMKLYFAQTLLHIGSS----HLAPPNDKQFLISPPAPPGVKQVEDATPV 126

```

Db 71 LHETQFRGKGLKLYPAQVOTPETDQKLTALPPOPAKQFLSPSSPPYGMKPIDAIPV 130
OY 127 INVDLLYAIASXSGPGEKYLHAATDPTSPVVVAVCESDENESEEMBRMRPKPKIIQ 186
Db 131 LNYDLYAAVAKLGPGEKYLHAGTSTPSVVVHVCDSDMEEDERK-----TSPKPKIIQ 185
OY 187 TRRPEYTP 194
Db 186 TRRPGILPP 193

RESULT 5
CCP2_HUMAN STANDARD; PRT; 197 AA.
ID CCP2_HUMAN
AC 014206;
RT 30-MAY-2000 (Rel. 39, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Calcipressin 2 (Thyroid hormone-responsive protein ZAKI-4) (Down
DE syndrome candidate region 1-like 1) (Myocyte-enriched calcineurin
DE interacting protein 2) (MCIP2).
GN DSCR1L1 OR ZAKI4.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=fibroblast; PubMed=8662924;
RX MEDLINE=96278928; Pubmed=8662924;
RA Miyazaki T., Kanou Y., Murata Y., Ohmori S., Niwa T., Maeda K.,
RA Yamamura H., Seo H.,
RT "Molecular cloning of a novel thyroid hormone-responsive gene, ZAKI-4,
RT in human skin fibroblasts."
CC -1- FUNCTION: Inhibits calcineurin-dependent transcriptional responses
CC by binding to the catalytic domain of calcineurin A. Could play a
CC role during central nervous system development.
CC -1- TISSUE SPECIFICITY: Expressed in fibroblasts, heart, brain, liver,
CC and skeletal muscle but not in placenta, lung, kidney and
CC pancreas.
CC -1- INDUCTION: By thyroid hormone.
CC -1- SIMILARITY: BELONGS TO THE DSCR1 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC EMBL: D83407; BAA11911.1; ALT_INIT.
CC DR Genew; HGNC:3041; DSCR1L1.
CC DR MIM: 604876;
CC DR SEQUENCE 197 AA; 21955 MW; AF39735F6661CBE0 CRC64;
SQ

Query Match 59.0%; Score 615.5; DB 1; Length 197;
Best Local Similarity 63.3%; Pred. No. 6.7e-46;
Matches 119; Conservative 23; Mismatches 37; Indels 9; Gaps 2;

```

```

OY 187 TRRPEYTP 194
Db 186 TRRPGILPP 193

RESULT 6
CCP3_MOUSE STANDARD; PRT; 239 AA.
ID CCP3_MOUSE
AC 09JXKO; 09CX87;
RT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Calcipressin 3 (Down syndrome candidate region 1-like protein 2)
DE (Myocyte-enriched calcineurin interacting protein 3) (MCIP3).
GN DSCR1L2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BA1B/c; TISSUE=Brain;
RA Strippoli P., Petini M., Lenzi L., Carinci P., Zannotti M.;
RT "The murine DSCR1-like gene family."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 40-239 FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryonic head;
RX MEDLINE=21085660; Pubmed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aikawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Aikawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamahata I.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batilov S., Casavant T.,
RA Fleischnmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
RA Kuehl P., Lewis S., Marsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Stauber F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okada T., Furuno M., Aono H., Baldarelli R., Barn G.,
RA Blake J., Boffelli D., Bojunga N., Carinci P., de Bonaldo M.F.,
RA Brownstein M.J., But C., Fletcher C., Fujita M., Gariboldi M.,
RA Guerinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombauts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Welter C., Whitaker C., Wilting L.,
RA Wyszynski B., Yoshida K., Hasegawa Y., Kawai H., Kohzuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:665-690(2001).
CC -1- FUNCTION: Inhibits calcineurin-dependent transcriptional responses
CC by binding to the catalytic domain of calcineurin A. Could play a
CC role during central nervous system development (By similarity).
CC -1- SIMILARITY: BELONGS TO THE DSCR1 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC EMBL: AF237889; AAF62539.1;
CC DR EMBL; AK019377; BAB31687.1;
CC DR MGD; MGI:1858220; Dscr1l2.
CC DR SEQUENCE 239 AA; 27153 MW; 1B2687B47B8AD272 CRC64;
SQ

Query Match 57.0%; Score 594.5; DB 1; Length 239;
Best Local Similarity 60.4%; Pred. No. 5.3e-44;
Matches 122; Conservative 19; Mismatches 48; Indels 13; Gaps 3;

```

```

Db 40 MDLSLP---TSLFACSVHEAFVQEKERFEALFTLYDDQVTFQFKSFRRVINFSK 96
Qy 61 PLISAADARLRHLKHTFLGKEMKLYFAQTLHIG-----SSHAPPNDKQFLISPASPVG 116
Db 97 PEAAARARIELHESFGRKLYFAQVQSGEARDKSYLLPPQTKQFLISPASPVG 156
Qy 117 WKQVEDATPVINYDLYLAIKLGPGKGYELHAATDTPSVVHVHVCESDOENEEEMER 176
Db 157 WKQSEDAMPVINYDLYLCAVSLGPGKGYELHAGTSTPSTVSVHVHVCESETEED----- 210
Qy 177 MKRPKPKIITRRPEYTHLS 198
Db 211 TKNPQKQITQRRPEAPTAALS 232

RESULT 7
CCP3 HUMAN
ID _CCP3 HUMAN STANDARD; PRT; 241 AA.
AC O9UKA8; O9UKA7; O9NUC8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Calpriesin 3 (Down syndrome candidate region 1-like protein 2)
GN (Myocyte-enriched calcineurin interacting protein 3) (MCIP3).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Peripheral blood, and Placenta;
RX MEDLINE=2021370; PubMed=10756093;
RA Strippoli P., Lenzi L., Petrini M., Carinci P., Zannotti M.;
RT "A new gene family including DSCR1 (Down syndrome candidate) and
RT DSCR1-like 2: characterization from yeast to human and identification of
RL Genomics 64:252-263(2000).
RN [2]
RP SEQUENCE OF 66-241 FROM N.A.
RA Bagguley C.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Inhibits calcineurin-dependent transcriptional responses
CC by binding to the catalytic domain of calcineurin A. Could play a
CC role during central nervous system development (By similarity).
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (SHOWN HERE) and 2; are
CC produced by alternative splicing.
CC -1- TISSUE SPECIFICITY: Highest expression in heart, skeletal muscle
CC kidney, liver and peripheral blood leukocytes. Lower expression in
CC all other tissues.
CC -1- SIMILARITY: BELONGS TO THE DSCR1 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF176116; AAF01684.1; -
CC DR EMBL; AF176117; AAF01685.1; -
CC DR EMBL; AL034582; CAB72298.1; -
CC DR Genew; HGNC:13042; DSCR1L2.
CC DR MIM; 605860; -
CC KW Alternative splicing.
CC FT VARSPLOC 124 133 MISSING (IN ISOFORM 2).
CC SQ SEQUENCE 241 AA; 27492 MW; 91AB619F5E443FBD CRC64;

Query Match 56.8%; Score 592.5; DB 1; Length 241;
Best Local Similarity 61.5%; Pred. No. 86-44;
Matches 120; Conservative 19; Mismatches 43; Indels 13; Gaps 3;

```

```

Qy 1 MDRFDSYFSSLIACVANDDVFSSETRAKFSLFRTYDKDTTFQYFKSKFRVINFSN 60
Db 40 MDLSLP---TSLFACSVHEAFVQEKERFEALFTLYDDQVTFQFKSFRRVINFSK 96
Qy 61 PLISAADARLRHLKHTFLGKEMKLYFAQTLHIG-----SSHAPPNDKQFLISPASPVG 116
Db 97 PEAAARARIELHESFGRKLYFAQVQSGEARDKSYLLPPQTKQFLISPASPVG 156
Qy 117 WKQVEDATPVINYDLYLAIKLGPGKGYELHAATDTPSVVHVHVCESDOENEEEMER 176
Db 157 WKQSEDAMPVINYDLYLCAVSLGPGKGYELHAGTSTPSTVSVHVHVCESETEED----- 210
Qy 177 MKRPKPKIITRRPE 191
Db 211 TKNPQKQIATRRPD 225

RESULT 8
NLA DROME
ID _NLA DROME STANDARD; PRT; 292 AA.
AC O9XZL8; O9V391;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Nebula protein.
GN NLA OR CG6072;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA McCormack A.V., Goldberg M.L.;
RT "Gene required for elongation of meiosis I spindle in Drosophila
RT females.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.V., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,
RA Fostel C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Ke Z., Kennison J.A., Ketchum K.A.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Wei M.-H., Ibegwam C.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paciej J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

```



```

RN [1]
RN SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=95091838; PubMed=7999131;
RA Oka J., Matsumoto A., Hosokawa Y., Inoue S.;
RT "Molecular cloning of human cytosolic purine 5'-nucleotidase.";
RN Biochem. Biophys. Res. Commun. 205:917-922(1994).
RN [2]
RN SEQUENCE FROM N.A.
RC TISSUE=Lymph;
RA Straube R.;
RT Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MAY HAVE A CRITICAL ROLE IN THE MAINTENANCE OF A
CC CONSTANT COMPOSITION OF INTRACELLULAR PURINE/PYRIMIDINE
CC NUCLEOTIDES IN COOPERATION WITH OTHER NUCLEOTIDASES.
CC PREFERENTIALLY HYDROLYSES INOSINE 5'-PRIME-MONOPHOSPHATE (IMP) AND
CC OTHER PURINE NUCLEOTIDES.
CC -1- CATALYTIC ACTIVITY: A 5'-ribonucleotide + H(2)O = a ribonucleoside
CC + phosphate.
CC -1- ENZYME REGULATION: ALLOSTERICALLY ACTIVATED BY VARIOUS COMPOUNDS,
CC INCLUDING ATP.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- DISEASE: ABERRATION OF THIS ENZYME MAY BE RELATED TO THE URATE
CC PRODUCTION IN HYPERURICEMIA AND GOUT.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D38524; BA007529.1; -
DR EMBL; BC001595; AA001595.1; -
DR Genew; HGNC:8022; NT5C2.
DR MIM; 600417; -
KW Hydroxylase; Allosteric enzyme.
FT DOMAIN 549 561 ASP/GLU-RICH (ACIDIC).
SQ SEQUENCE 561 AA; 64969 MW; 4C27D762575BDEA2 CRC64;

Query Match 8.2%; Score 85.5; DB 1; Length 561;
Best Local Similarity 24.2%; Pred. No. 5.8;
Matches 47; Conservative 29; Mismatches 71; Indels 47; Gaps 10;

QY 11 SLLIACVANDVDFSESETRAKFSLPR-----TYDKOTTP-----QYFKSPK 52
DB 386 SLLFEELQSLDIFLAELVYHLDSSNERPDISTORRIKVTHTDMDCYGMNGSLFRSGS 445
QY 53 RVRINFSNPLSAADARLRLHKTFLGKEMKLYFAOTLHIGSHLAPPNP-----DKQ 104
DB 446 RQTLRPSQVMRYAD-----LYASFIN-----LVVYPSYLPRAHVLVMPHSTVEHTVDIN 498
QY 105 FLISPPAS---PVGKQVEDATPVINVDLLVIAISKLGKGEKELMAADPPPSVVHVC 161
DB 499 EMESPLATNRTSVDFKD-----TDYKHQOLTRISISIKPNNLFPPL-----APOELTH-C 547
QY 162 ESDQNEEEEEEME 175
DB 548 HDDEDEEEEEEEE 561

```

```

OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=5288C;
RX MEDLINE=94378719; PubMed=8091858;
RA Vandendol W., Bolle P.-A., Dion C., Portetelle D., Hliger F.;
RT "Sequencing and analysis of a 20.5 kb DNA segment located on the left
RT arm of yeast chromosome XI.";
RL Yeast 10:525-S33(1994).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z26878; CA81519.1; -
DR EMBL; Z28168; CA82010.1; -
DR PIR; S37998; S37998.
DR PIR; S38413; S38413.
DR PIR; S44593; S44593.
DR SGD; S0001651; YKL168C.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE; PS00108; PROTEIN_KINASE_ST_1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW Hypothetical protein; Transferase; Serine/threonine-protein kinase;
KW ATP-binding.
FT DOMAIN 422 722
FT NP_BIND 428 436 ATP (BY SIMILARITY).
FT BINDING 465 465 ATP (BY SIMILARITY).
FT ACT_SITE 573 573 BY SIMILARITY.
SQ SEQUENCE 734 AA; 83656 MW; 896D26B060A917D CRC64;

Query Match 7.9%; Score 82.5; DB 1; Length 734;
Best Local Similarity 21.4%; Pred. No. 15;
Matches 42; Conservative 29; Mismatches 64; Indels 61; Gaps 8;

QY 11 SLLIACVANDVDFSESETRAKFSLPR-----TYDKOTTP-----FOYFKSPK 53
DB 155 SSVSCDSNSNGTSSSDQWAMDSDLDDSDNDLTPYRGSNKDILSKDRAPVNYIDYXK 214
QY 54 VRI---NFSNPLSAADARLRLHKTFLGKEMKLYFAOTLHIGSHLAPPNPDKQFLIS 109
DB 215 XALRATSYPLPLPS-----KQYNERLYTRSH-----PDESLSL 252
QY 110 P--ASPPV-----GKQVEDATPVINVDLLVIAISKLGKGEKELH---AATDTPSPV 157
DB 253 PRFAQADVQCIIIEONGFVYEDGSHENIKLSGVIAKLEKNSLPRAHQGSLSRRLGIT 312
QY 158 V-----HVCSDDEN 167
DB 313 LSGLFKHKHKNECDIEN 328

```

```

RESULT 13
K08_YEAST STANDARD; PRT; 734 AA.
ID_K08_YEAST
AC P36004;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable serine/threonine-protein kinase YKL168C (EC 2.7.1.1.-).
GN YKL168C OR YKL632.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

```

```

RESULT 14
CR42_BACSH STANDARD; PRT; 370 AA.
ID_CR42_BACSH
AC P0516;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE 41.9 kDa insecticidal toxin (Strain 2297).
GN B1N.
OS Bacillus sphaericus.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

```

```

OX NCBI_TaxID=1421;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88234018; PubMed=3375083;
RA Hindley J., Berry C.;
RT "Bacillus sphaericus strain 2297: nucleotide sequence of 41.9 kDa
RL toxin gene";
RN Nucleic Acids Res. 16:4168-4168(1988).
RP [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=88197989; PubMed=3360740;
RA Baumann L., Broadwell A.H., Baumann P.;
RT "Sequence analysis of the mosquitoicidal toxin genes encoding 51.4-
RN and 41.9-kilodalton proteins from Bacillus sphaericus 2362 and
RT 2297.";
RL J. Bacteriol. 170:2045-2050(1988).
CC CC -1- FUNCTION: THIS TOXIN IS ACTIVE AGAINST CULEX AND ANOPHELES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X07025; CAA30074.1; -
DR PIR; S01856; S01856.
DR PIR; D28211; D28211.
KW Toxin; Sporulation.
SQ SEQUENCE 370 AA; 41917 MW; 545D66E9505911D9 CRC64;

Query Match 7.9%; Score 82; DB 1; Length 370;
Best Local Similarity 21.1%; Pred. No. 6.9;
Matches 40; Conservative 27; Mismatches 79; Indels 44; Gaps 7;

QY 1 MDRFDSYNFSSLIACVANDVFE--SETRAFESLFRFYD----- 40
Db 21 MDFNSYFPCIHAPSAPNGDITEICSRNNQYFFFPDDGRVIANRHNSVFTGEA 80
QY 41 -----KDTTFQYKSFRRVINF-----SNPLSADARLRHLKTEFLGKEMKLYPAQT 88
Db 81 TSVVVDIYTGSPLOQFFREFRKTSTYLAIQNPESATDVRALEPNSHEL--PSRLFTNN 138
QY 89 LHGTSSHLAPNFDKQELISPPASPVGWQVEDATPVINYDLYAISKLGPGKEYELHA 148
Db 139 IENNSNLI--SNKEQIYLTPLSLPE---NEQYKTPVLS-----GDDIGPNQSEKSI 188
QY 149 ATDPTSVVV 158
Db 189 GSTLIPCIMV 198

RESULT 15
SEPA EMENI
ID SEPA EMENI STANDARD; PRT; 1790 AA.
AC P78621; Q00760;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cytokinesis protein sepa (Phi/2 protein) (Forced expression
DE inhibition of growth A).
GN SEPA OR FIGA.
OS Emericella nidulans (Aspergillus nidulans).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; Emericella.
OX NCBI_TaxID=5072;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97361829; PubMed=9218790;
RA Harris S.D., Hamer L., Sharpless K.E., Hamer J.E.;
RT "The Aspergillus nidulans sepa gene encodes an Phi/2 protein involved
RT in cytokinesis and the maintenance of cellular polarity.";
```

```

RL EMBO J. 16:3474-3483(1997).
RN [2]
RP REVISIONS TO 143-153; 207; 1071-1109 AND 1644.
RA Hamer L., Harris S.D., Sharpless K.E., Hamer J.E.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 559-1790 FROM N.A., AND FUNCTION.
RX STRAIN=RGSC 26;
RX MEDLINE=95229045; PubMed=7713416;
RA Marthou J.F., Adams T.H.;
RT "Identification of developmental regulatory genes in Aspergillus
RN nidulans by overexpression.";
RL Genetics 139:537-547(1995).
CC CC -1- FUNCTION: INVOLVED IN CYTOKINESIS. OVEREXPRESSION RESULTS IN
CC GROWTH INHIBITION.
CC -1- DOMAIN: DRFS ARE REGULATED BY INTRAMOLECULAR GBD-DAD BINDING WHERE
CC RHO-GTP ACTIVATES THE DRFS BY DISRUPTING THE GBD-DAD INTERACTION
CC (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 GTPASE-BINDING DOMAIN (GBD).
CC -1- SIMILARITY: CONTAINS 1 FORMIN HOMOLGY 1 (FH1) DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 FORMIN HOMOLGY 2 (FH2) DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 FORMIN HOMOLGY 3 (FH3) DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 DRF AUTOREGULATORY DOMAIN (DAD).
CC -1- SIMILARITY: BELONGS TO THE FORMIN HOMOLGY FAMILY. BNI1
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U83658; AAB63335.3; -
DR EMBL; L36341; AAA33306.1; ALT_SEQ.
DR InterPro; IPR003104; FH2.
DR Pfam; PF02181; FH2; 1.
DR SMART; SM00498; FH2; 1.
KW Cell division; Coiled coil.
FT DOMAIN 258 486 GBD.
FT DOMAIN 376 718 FH3.
FT DOMAIN 724 811 COILED COIL (POTENTIAL).
FT DOMAIN 955 1136 FH1 (PRO-RICH).
FT DOMAIN 1141 1658 FH2.
FT DOMAIN 1435 1566 COILED COIL (POTENTIAL).
FT DOMAIN 1586 1600 DAD.
FT DOMAIN 1608 1611 ARG/LYS-RICH (BASIC).
FT CONFLICT 1071 1109 PPPPPPGGFGGPPPPPPPPPGGFGGPPPPPPPPPPPGGFGG
FT -> LPSPSSTGFWMTTPAASPTWFWRSTGAAAAASYWCI
FT C (IN REF. 3).
FT CONFLICT 1476 1476 D -> V (IN REF. 3).
FT CONFLICT 1504 1504 V -> L (IN REF. 3).
SQ SEQUENCE 1790 AA; 197355 MW; 192136DE2EF2A75B CRC64;

Query Match 7.8%; Score 81.5; DB 1; Length 1790;
Best Local Similarity 29.9%; Pred. No. 54;
Matches 26; Conservative 11; Mismatches 45; Indels 5; Gaps 2;

QY 98 PNPDKQELISPPASPP-----VGWQVEDATPVINYDLYAISKLGPGKEYE-LHAATDP 152
Db 1101 PPPPGGAGFVPPPPPPPGTGVIGGWRANYLASQAPSHAIPVWSSIRPKKALKALHWDKVD 1160
QY 153 TPSVVVHVCSQDNEEEEEEMERMKR 179
Db 1161 TPQVTVWATHGTTPOEKEEYVELAKR 1187

Search completed: December 11, 2002, 11:39:41
Job time : 9.3457 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2002, 14:10:15 ; Search time 97.8236 Seconds
(without alignments)
13789.580 Million cell updates/sec

Title: US-09-782-953-1
Perfect score: 599
Sequence: 1 gaggtgcaaggacacctcca.....gaggaggaagagagagat 599

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_101002:*

- 1: /SID27/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
- 2: /SID27/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
- 3: /SID27/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
- 4: /SID27/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
- 5: /SID27/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
- 6: /SID27/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
- 7: /SID27/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
- 8: /SID27/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
- 9: /SID27/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
- 10: /SID27/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
- 11: /SID27/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
- 12: /SID27/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
- 13: /SID27/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
- 14: /SID27/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
- 15: /SID27/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
- 16: /SID27/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
- 17: /SID27/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
- 18: /SID27/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
- 19: /SID27/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
- 20: /SID27/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
- 21: /SID27/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
- 22: /SID27/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
- 23: /SID27/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
- 24: /SID27/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	599	100.0	599	24	Human MCIP associa
2	504.4	84.2	597	24	Mouse MCIP associa
3	426.2	71.2	597	24	Human MCIP associa
4	413.8	69.1	2331	24	Human MCIP associa
5	404.8	67.6	2348	24	Human DNA sequence
6	357.8	59.7	2358	21	Lung cancer associ
7	349.2	58.3	2212	24	Human MCIP associa
8	341.8	57.1	2173	24	Gene #3341 used to
9	340.2	56.8	2174	20	Human DSCR1 coding

10	264.4	44.1	442	21	AAC01774	Human secreted pro
11	202.2	33.8	3184	24	AAD30157	Human MCIP associa
12	202.2	33.8	3184	24	ABL61768	Colon adenocarcino
13	202.2	33.8	3184	24	ABL65187	Lung cancer relate
14	197.6	33.0	934	22	AAF25338	Nucleotide sequenc
15	197.6	33.0	3159	24	ABA91457	Human Down syndrom
16	191	31.9	594	24	AAD30154	Mouse MCIP associa
17	189.8	31.7	720	24	AAD30159	Human MCIP associa
18	188.4	31.5	828	24	AAD30158	Human MCIP associa
19	181.8	30.4	615	24	ABA91463	Rat Down syndrome
20	153	25.5	1021	24	ABA91458	Rat Down syndrome
21	151.6	25.3	640	24	ABQ59719	Human colon cancer
22	144.6	24.1	412	22	ABA56172	Human foetal liver
23	144.6	24.1	412	22	AK04364	Human brain expres
24	144.6	24.1	412	22	AA114449	Probe #4382 for ge
25	144.6	24.1	412	22	AA135821	Probe #4507 used t
26	144.6	24.1	412	22	AA104272	Probe #4263 used t
27	144.6	24.1	412	24	ABSO4419	Human genome-deriv
28	144.6	24.1	446	22	ABA43406	Human breast cell
29	144.6	24.1	446	22	ABA53852	Human foetal liver
30	144.6	24.1	446	22	ABA23598	Probe #2064 for ge
31	144.6	24.1	446	22	AAK02113	Human brain expres
32	144.6	24.1	446	22	AAK27563	Human bone marrow
33	144.6	24.1	446	22	AA112148	Probe #2081 for ge
34	144.6	24.1	446	22	AA133493	Probe #2179 used t
35	144.6	24.1	446	22	AA102061	Probe #2052 used t
36	144.6	24.1	446	24	ABS02041	Human genome-deriv
37	91.8	15.3	111	22	ABA68757	Human foetal liver
38	91.8	15.3	111	22	AAK17102	Human brain expres
39	91.8	15.3	111	22	AA123650	Probe #13583 for g
40	91.8	15.3	111	22	AA148960	Probe #17646 used
41	91.8	15.3	111	22	AA109265	Probe #9256 used t
42	91.8	15.3	111	22	ABS16940	Human genome-deriv
43	77.4	12.9	123	22	ABA47285	Human breast cell
44	77.4	12.9	123	22	ABA65170	Human foetal liver
45	77.4	12.9	123	22	ABA32272	Probe #10738 for g

ALIGNMENTS

RESULT 1
AAD30151
ID AAD30151 standard; DNA; 599 BP.
XX
AC AAD30151;
XX

DT 17-MAY-2002 (first entry)

DE Human MCIP associated DNA #1.

DE Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;
KW heart failure; cardiomyopathy; heart disease; human; ds.
XX

OS Homo sapiens.

PN WO200204491-A2.

PD 17-JAN-2002.

PF 06-JUL-2001; 2001WO-US21662.

PR 07-JUL-2000; 2000US-216601P.

PR 13-FEB-2001; 2001US-0782953.

PA (TEXA) UNIV TEXAS SYSTEM.

PA (WILL/) WILLIAMS S R.

PA (ROTH/) ROTHERMEL B.

PI Williams SR, Rothermel B;

XX WPI; 2002-179698/23.

Oy	1	GAGGTGCAAAAGAACCTCTCACCTTGGGGTTGATCGAAGAGACGAGTGTGTTAAAGCGT	60
Db	1	GAGGTGCAAAAGAACCTCTCACCTTGGGGTTGATCGAAGAGACGAGTGTGTTAAAGCGT	60
Oy	61	CTGCCCCGTGAAAAAGCAGATGATTTTGGAGCTTAGCTACATTTTAGTCCCTGAT	120
Db	61	CTGCCCCGTGAAAAAGCAGATGATTTTGGAGCTTAGCTACATTTTAGTCCCTGAT	120
Oy	121	TGCTTGTTGTGCGAAACGATGATGTCTTCAGGGAAAGTGAACGAGGGCCAAATTTGAATC	180
Db	121	TGCTTGTTGTGCGAAACGATGATGTCTTCAGGGAAAGTGAACGAGGGCCAAATTTGAATC	180
Oy	181	CCTCTTCAGAACATATGACAGAGACACCCTTCAGATATTTTAAAGCTTCAAAAGTGT	240
Db	181	CCTCTTCAGAACATATGACAGAGACACCCTTCAGATATTTTAAAGCTTCAAAAGTGT	240
Oy	241	CCGAGTAAACTTCAGCAACCCCTTATCTGACGCGATGCCAGGTGCGGCTGCACAAAGAC	300
Db	241	CCGAGTAAACTTCAGCAACCCCTTATCTGACGCGATGCCAGGTGCGGCTGCACAAAGAC	300
Oy	301	CGAGTTCTGSGGGAAGAAATGAAATGATTTTGTCTCAGACTTAAACATPAGGAAGTTC	360
Db	301	CGAGTTCTGSGGGAAGAAATGAAATGATTTTGTCTCAGACTTAAACATPAGGAAGTTC	360
Oy	361	AACCTGGGCTCGGCAATCCGGAACAAAGTTCCATCTCCCTCGGGGCTCTCCCTCC	420
Db	361	AACCTGGGCTCGGCAATCCGGAACAAAGTTCCATCTCCCTCGGGGCTCTCCCTCC	420
Oy	421	GTTGGCTGGAACAAGTAGAAGATCCACCCCGCTCATAAATTAATTCATCTTTATATGCC	480
Db	421	GTTGGCTGGAACAAGTAGAAGATCCACCCCGCTCATAAATTAATTCATCTTTATATGCC	480
Oy	481	ATCTCCAACTGGGGCCAGAGAGAAAGTATGATGTCATGCGACGACAGACACCACTCCC	540
Db	481	ATCTCCAACTGGGGCCAGAGAGAAAGTATGATGTCATGCGACGACAGACACCACTCCC	540
Oy	541	AGTGTGTGTGTTCCAGTGTGTGTGATGATCCAAAGAAATGAGAGAGAAAGGAAGAT	599
Db	541	AGTGTGTGTGTTCCAGTGTGTGTGATGATCCAAAGAAATGAGAGAGAAAGGAAGAT	599

RESULT 2	
AAD30153	
ID	AAD30153 standard; DNA; 597 BP
XX	
AC	AAD30153;
XX	

DB 124 GACACCACCTTCAGTATTTAAGAGCCTCCATACGTCGCCGAATCATCTCTGCATCCCC 107

QY 263 TTATCTGAGCCGATGCGAGCTGGCTGCAAGACCGAGTTCTTGGGGAAGGAATG 322
 Db 184 TTATCTGAGCCGATGCGAGCTGGCTGCAAGACCGAGTTCTTGGGGAAGGAATG 243
 QY 323 AAGTTGTATTTTGTCTGAGACTTTACACATAGGAAGTTTACACCTGGCTCCG-CCAAATCCC 381
 Db 244 AAGTTGTATTTTGTCTGAGACTTTACACATAGGAAGTTTACACCTGGCTCCG-CCAAATCCC 303
 QY 382 GACAAACAGTTTCTCATCTCCCTCCGGCTCTCTCTCCCTGTTGGCTGGAACAAGTAGAA 441
 Db 304 GACAAACAGTTTCTCATCTCCCTCCGGCTCTCTCTCCCTGTTGGCTGGAACAAGTAGAA 363
 QY 442 GATCCACCCCGTCATAAATACGATCTTTTATATGCGATCTCCAACTGGGCGCAGGA 501
 Db 364 GATCCACCCCGTCATAAATACGATCTTTTATATGCGATCTCCAACTGGGCGCAGGA 423
 QY 502 GAGAAGTATGAATGTCATGTCAGGAGACACACCACTCCAGTGTGGTGTCCACGTGTGT 561
 Db 424 GAGAAGTATGAATGTCATGTCAGGAGACACACCACTCCAGTGTGGTGTCCACGTGTGT 483
 QY 562 GAGAGTACCAAGAGATGAGGAGGAAGAGAGAGAT 599
 Db 484 GAGAGTACCAAGAGATGAGGAGGAAGAGAGAGAT 521

RESULT 3
 AAD30152
 ID AAD30152 standard; DNA; 597 BP.
 AC AAD30152;
 DT 17-MAY-2002 (first entry)
 DE Human MCIP associated DNA #2.
 KW Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;
 KW heart failure; cardiomyopathy; heart disease; human; gene; ds.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 CDS 1..597
 FT /*tag= a
 FT /product= "Human MCIP associated protein #2"
 PN WO200204491-A2.
 XX 17-JAN-2002.
 XX 06-JUL-2001; 2001WO-US21662.
 XX 07-JUL-2000; 2000US-216601P.
 PR 13-FEB-2001; 2001US-0782953.
 XX (TEXA) UNIV TEXAS SYSTEM.
 PA (WILL/) WILLIAMS S R.
 PA (ROTH/) ROTHERMEL B.
 XX Williams SR, Rothermel B;
 PI WPI; 2002-179698/23.
 DR P-PSDB; AAE18911.
 XX Screening for modulators of muscle calcineurin interacting protein
 PT (MCIP) binding, expression or phosphorylation, useful for treating
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,
 PT calcineurin and a test compound .
 XX Claim 95; Page 147-148; 174pp; English.
 PS The invention relates to muscle calcineurin interacting proteins (MCIPs)
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical
 CC complex with the catalytic subunit of calcineurin and increased levels

CC of MCIPs correspond to a reduced ability of calcineurin to stimulate
 CC transcription of certain target genes. The invention also relates to
 CC methods for identifying modulators of MCIP binding, expression or
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin
 CC may be used for treating cardiac hypertrophy and heart failure.
 CC Antibodies to MCIP can be used in characterising the MCIP content of
 CC healthy and diseased tissues and subsequently for determining the
 CC presence or absence of cardiomyopathy or as predictor of heart disease.
 CC The present sequence is human MCIP associated DNA.
 XX

SQ Sequence 597 BP; 160 A; 170 C; 151 G; 116 T; 0 other;

Query Match 71.2%; Score 426.2; DB 24; Length 597;
 Best Local Similarity 93.1%; Pred. No. 9e-128;
 Matches 457; Conservative 0; Mismatches 33; Indels 1; Gaps 1;

QY 110 AGCTCCCTGATGCTTGTGTGGCAACGATGATGCTTCAGCGAAAGTGAGACCGGGCC 169
 Db 31 AGCGCCACCATCGCTGCCACCTGGACCGCGCTGTTCTGGAGCGGCTGTGCCGGGCC 90
 QY 170 AAATTTGAATCCCTCTTCAGAACATATGACAAAGGACACCACTTCCAGTATTTTAAAGC 229
 Db 91 AAATTTGAATCCCTCTTCAGAACATATGACAAAGGACACCACTTCCAGTATTTTAAAGC 150
 QY 230 TTCAAACGTGTCGGATAAACTTCAGCAACCCCTTATCTGCAGCCGATGCCAGGCTGCGG 289
 Db 151 TTCAAACGTGTCGGATAAACTTCAGCAACCCCTTATCTGCAGCCGATGCCAGGCTGCGG 210
 QY 290 CTGCACAGACCGAGTTCTTGGGGAAGGAATGAAGTTGTATTTTGTCTCAGACTTTACAC 349
 Db 211 CTGCACAGACCGAGTTCTTGGGGAAGGAATGAAGTTGTATTTTGTCTCAGACTTTACAC 270
 QY 350 ATAGGAAGTTACACCTGGCTCCG-CCAATCCGACAAACAGTTCTCTCATCTCCCTCCG 408
 Db 271 ATAGGAAGTTACACCTGGCTCCG-CCAATCCGACAAACAGTTCTCTCATCTCCCTCCG 330
 QY 409 GCCTCTCTCCCGTTGGCTGGAAACAAGTAGAAGATGCCACCCCGTCAATAAATTACGAT 468
 Db 331 GCCTCTCTCCCGTTGGCTGGAAACAAGTAGAAGATGCCACCCCGTCAATAAATTACGAT 390
 QY 469 CTTTATATGCGATCTCCAAAGCTGGGCGCAGGAGAAAGTATGAATCTGCATGCGCGACA 528
 Db 391 CTTTATATGCGATCTCCAAAGCTGGGCGCAGGAGAAAGTATGAATCTGCATGCGCGACA 450
 QY 529 GACACCACTCCCACTGGTGGTCCACCTGTGTGAGAGTACCAGAGAAATGAGGAGGAA 588
 Db 451 GACCCCACTCCCACTGGTGGTCCACCTGTGTGAGAGTACCAGAGAAATGAGGAGGAA 510
 QY 589 GAGGAAGAGAT 599
 Db 511 GAGGAAGAGAT 521

RESULT 4
 AAD30155
 ID AAD30155 standard; DNA; 2331 BP.
 XX AAD30155;
 AC AAD30155;
 XX 17-MAY-2002 (first entry)
 DE Human MCIP associated DNA #1.

XX Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;
 KW heart failure; cardiomyopathy; heart disease; human; gene; ds.
 XX Homo sapiens.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 CDS 144..737
 FT /*tag= a
 FT /product= "Human MCIP associated protein #1"

PN WO200204491-A2.
 XX 17-JAN-2002.
 PD 06-JUL-2001; 2001WO-US21662.
 XX 07-JUL-2000; 2000US-216601P.
 PR 13-FEB-2001; 2001US-0782953.
 XX (TEXA.) UNIV TEXAS SYSTEM.
 PA (MILL.) WILLIAMS S R.
 PA (ROTH/) ROTHERMEL B.
 PI Williams SR, Rothermel B;
 XX WPI; 2002-179698/23.
 DR P-PSDB; AAE18914.
 XX Screening for modulators of muscle calcineurin interacting protein
 PT (MCIP) binding, expression or phosphorylation, useful for treating
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,
 PT calcineurin and a test compound -
 XX
 XX Claim 72; Page 155-157; 174pp; English.
 PS
 XX The invention relates to muscle calcineurin interacting proteins (MCIPs)
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical
 CC complex with the catalytic subunit of calcineurin and increased levels
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate
 CC transcription of certain target genes. The invention also relates to
 CC methods for identifying modulators of MCIP binding; expression or
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin
 CC may be used for treating cardiac hypertrophy and heart failure.
 CC Antibodies to MCIP can be used in characterizing the MCIP content of
 CC healthy and diseased tissues and subsequently for determining the
 CC presence or absence of cardiomyopathy or as predictor of heart disease.
 CC The present sequence is human MCIP associated DNA.
 CC Note: This sequence has been described as a promoter in claim 72 of
 CC the specification, however the sequence seems to be a polynucleotide
 CC encoding a MCIP associated protein.
 CC
 XX Sequence 2331 BP; 630 A; 470 C; 547 G; 684 T; 0 other;
 SO
 Query Match 69.1%; Score 413.8; DB 24; Length 2331;
 Best Local Similarity 85.1%; Pred. No. 1.9e-123;
 Matches 474; Conservative 0; Mismatches 82; Indels 1; Gaps 1;
 QY 39 GAGCGAGTCGTTGCTTAAGCGTCTGCCCCCGTGAAGGAGGATGATTTAGGAGACTTTA 98
 DB 103 GACTGGCGTGGTGTGTAGCCGCTTCACTGTAGAGAGGAGGATGATTTAGAACTTTA 162
 QY 99 GCTACATTTTAACTCCCTGATGCTTGTGTGGCAACGATGATGCTTCCAGGAAAGTG 158
 DB 163 ACTACAGTTTAACTCCCTGATGCTTGTGTGGCAACGATGATGCTTCCAGGAAAGTG 222
 QY 159 AGACAGGAGGCAATTTGAATCCCTCTTCAAGAACATGTGACAGAGACACCTTCCAGT 218
 DB 223 AAACAGGAGGCAATTTGAATCCCTCTTCAAGAGATGACAGAGACATCACCTTTCAGT 282
 QY 219 ATTTTAAGAGCTTCAAGCGTCTCCGATTAATCTGAGCAACCCCTTATCGAGCGGATG 278
 DB 283 ATTTTAAGAGCTTCAAGCGATGAGTAAGATTAATCTGAGCAACCCCTTATCGAGCGATG 342
 QY 279 CCAAGGCTGGCGGTGACAGAACCGAGTCTCTGGGGAAGAAATGAATGTTATTTGCTC 338
 DB 343 CCAAGGCTGGCGGTGACAGAACCGAGTCTCTGGGGAAGAAATGAATGTTATTTGCTC 402
 QY 339 AGACTTTACACATAGAGAGTTCACACCTGCTCCGCC-AATCCGACAAACAGTTCTCTCA 397
 DB 403 AGACCTTACACATAGAGAGTTCACACCTGCTCCGCCAAATTCAGACAAACAGTTCTCTCA 462
 QY 398 TCTTCCTCTGGGCTCTCTCTCCCTTGGGGAAGAAAGATGAGAGGACACCCCGCTCA 457

DB 463 TCTCCCTCCCGCCTCTCCGCGAGTGGATGAAACAGTGAAGATGCCACCCAGTCA 522
 QY 458 TAAATTACGATCTTTATATATCCCATCTTCCAAAGCTGGGCGCCAGAGAGAGATATGACTGC 517
 DB 523 TAACTATGATCTCTTATATATCCCATCTTCCAAAGCTGGGCGCCAGAGAGAGATATGACTGC 582
 QY 518 ATGAGCGACAGACACCACTCCGATGGTGTGTCACGCTGTGTGAGAGTACCAAGAGA 577
 DB 583 ACGGAGCGACTGACACCACTCCGATGGTGTGTCACGCTGTGTGAGAGTATGATCAAGAGA 642
 QY 578 ATGAGAGAGAGAGAGAA 594
 DB 643 AGAGAGAGAGAGAGAA 659
 RESULT 5
 AAS94805
 ID AAS94805 standard; DNA; 2348 BP.
 AC AAS94805;
 DT 14-FEB-2002 (first entry)
 XX
 DE Human DNA sequence #60 expressed during foam cell differentiation.
 KM Human; foam cell differentiation; atherosclerosis; cerebral stroke;
 KM cardiovascular disorder; coronary artery disease; gene therapy; db.
 XX
 OS Homo sapiens.
 PN WO200177389-A2.
 PD 18-OCT-2001.
 PF 04-APR-2001; 2001WO-US11128.
 PR 05-APR-2000; 2000US-195106P.
 PA (INCY-) INCYTE GENOMICS INC.
 PI Shiffman D, Somogyi R, Lawn R, Sellhammer JF, Porter GJ, Mikita T;
 PI Tai J;
 DR WPI; 2002-010925/01.
 XX Composition useful for diagnosis of conditions, disorders or diseases
 PT associated with atherosclerosis, comprises several polynucleotides that
 PT are differentially expressed in foam cell development -
 XX
 PS Claim 1; Page 112-113; 315pp; English.
 XX
 CC The present invention relates to the isolation of human polynucleotide
 CC sequences that are differentially expressed during foam cell
 CC differentiation. The polynucleotide sequences of the invention or a
 CC composition comprising these polynucleotides are useful as a high
 CC throughput method for detecting altered expression of one or more
 CC polynucleotides in a sample. The polynucleotides can be used in the
 CC diagnosis of disorders associated with foam cell development such as
 CC atherosclerosis, cerebral stroke, and cardiovascular disorders such as
 CC coronary artery disease. The polynucleotide sequences can also be used
 CC as PCR primers and probes. The polynucleotides of the invention are also
 CC useful in gene therapy. AAS9476-AAS95021 represent the human
 CC polynucleotide sequences of the invention which are differentially
 CC expressed during foam cell differentiation.
 CC
 XX Sequence 2348 BP; 638 A; 473 C; 550 G; 687 T; 0 other;
 SO
 Query Match 67.6%; Score 404.8; DB 24; Length 2348;
 Best Local Similarity 84.1%; Pred. No. 1.6e-120;
 Matches 469; Conservative 0; Mismatches 87; Indels 2; Gaps 1;
 QY 39 GAGCGAGTCGTTGCTTAAGCGTCTGCCCCCGTGAAGGAGGATGATTTAGGAGACTTTA 98

Db 110 GACTCGTGGGCTCTAGCGCTTTCACGTGTAAGAAAGCAAGATGCAATTTTAAACATTTA 169
Qy 99 GGTACAAATTTAGTCCCTGATTTGCTGTGGCAAAAGCATGATGCTTTACGCAAGAGTG 158
Db 170 ACTACAGTTTTAGCTCCCTGATTTGCTGTGGCAAAAGCATGATGCTTTACGCAAGAGTG 229
Qy 159 AGACAGGCGCAAAATTTCAATCCCTTTCAGAACATATGACAAAGACACACCTTCAGT 218
Db 230 AAACAGGCGCAAAATTTGAGTCCCTTTCAGAACATATGACAAAGACATCACCTTTCACT 289
Qy 219 ATTTTAAGAGCTTCAAAAGCTGTCCGGATAAACTTCAGCAACCCCTTATCTGCAGCCGATG 278
Db 290 ATTTTAAGAGCTTCAAAAGCTGAGATTAACCTTCAGCAACCCCTTTCGCGACAGATG 349
Qy 279 CCAGCTCGGCTGCACAAAGCCAGTTCTTGGGAAAGGAATGAAGTTATTTTTCCTC 338
Db 350 CCAGCTCCAGCTGCATAAGACTGAGTTTCTGGGAAAGGAATGAAGTTATTTTTCCTC 409
Qy 339 AGACTTTACATAGGAAGTTTCACACCTGG--CTCCGCCAATCCCGAACAACAGTTCCTC 396
Db 410 AGACTTTACATAGGAAGTTTCACACCTGGCTCCGCCAATCCCGAACAACAGTTCCTC 469
Qy 397 ATCTCCCTCCGGCTCTCTCCCTGGTGGTGGAAACAAGTAGAAGATGCCACCCCGTC 456
Db 470 ATCTCCCTCCGGCTCTCCGCCAGTGGATGGAACAAGTGGACCCCGTC 529
Qy 457 ATAAATTACGATCTTTATATGATCCATCTCAAGCTGGGCGCAGAGAGAGTATGAAGT 516
Db 530 ATAAATTACGATCTTTATATGATCCATCTCAAGCTGGGCGCAGAGAGAGTATGAAGT 589
Qy 517 CATCGAGCAGACACACCTCTCCAGTGTGGTGGTGGAAACAAGTAGAAGTGGACCAAGAG 576
Db 590 CACGACGAGTACACCACTCCAGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 649
Qy 577 AATGAGGAGGAAGGAA 594
Db 650 AAGGAGGAAGGAGGAA 667

RESULT 6
AAF18328
ID AAF18328 standard; DNA; 2358 BP.
AC AAF18328;
DT 14-MAR-2001 (first entry)
DE Lung cancer associated polynucleotide sequence SEQ ID 347.
KW Human; lung cancer associated protein; neuroprotective; cytostatic;
KW cardioactive; immunomodulatory; muscular active; vulnerary;
KW gastrointestinal; nephrotropic; antiinfective; gynecological;
KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
KW proliferative disorder; wound healing; infectious disease; ds.
OS Homo sapiens.
XX
XX WO200055180-A2.
XX
XX 21-SEP-2000.
XX
XX 08-MAR-2000; 2000WO-US05918.
XX
XX 12-MAR-1999; 99US-0124270.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX (ROSE/) ROSEN C A.
XX
XX Ruben SM;
XX
XX WPI; 2000-587514/55.
XX P-PSDB; AAB58452.
XX

PT Lung cancer associated gene sequences, referred to as lung cancer
PT antigens, useful for treatment, prevention, and diagnosis of disorders
PT such as lung cancer -
XX
XX
PS Claim 1; Page 805-806; 1425pp; English.
XX
XX Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer
CC associated proteins represented in AAB58106 - AAB58548. Lung cancer
CC associated proteins and polynucleotide sequences, their agonists, and
CC antagonists may have neuroprotective, cytostatic, cardioactive, and
CC immunomodulatory; muscular active general; vulnerary; gastrointestinal
CC general; nephrotropic; antiinfective; gynecological; or antibacterial
CC activity. The invention also includes antibodies specific for the
CC protein or polynucleotide sequences. The lung cancer associated
CC polynucleotide sequences may be used for detection of lung cancer,
CC chromosome identification, as chromosome markers, and for numerous other
CC diagnostic or research purposes. The proteins may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders. The proteins may also be used in the treatment of wounds and
CC infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and
CC peptide AAB58549 are used in the course of the invention for the
CC identification and characterisation of the polynucleotide and protein
CC sequences.
XX
SQ Sequence 2358 BP; 641 A; 482 C; 584 G; 647 T; 4 other;
Query Match 59.7%; Score 357.8; DB 21; Length 2358;
Best Local Similarity 80.5%; Pred. No. 2.9e-105;
Matches 416; Conservative 2; Mismatches 99; Indels 0; Gaps 0;
Qy 78 AGAATGATTTTGGGACCTTTAGCTACAAATTTAGCTCCCTGATTTGTTGTGGCAAAAG 137
Db 112 AGATGGAGGAGTGAGCTCGAGGACCTGCCAGCGCCACCATCCCTGTCCCTGGACC 171
Qy 138 ATGATGTTTTCAGGAAAGTGAGACACAGGCGCAAAATTTGAATCCCTCTTCAAGACATATG 197
Db 172 CGCGGTGTTCTGGACGGCTGTGCGGGCCAAATTTGAGTCCCTCTTTAGGACGATG 231
Qy 198 ACAAGGACACACCTTCCAGTATTTTAAAGCTTCAAAAGTGTCCGGATAAACTTCAGCA 257
Db 232 ACAAGGACATCACCTTTTCAGTATTTTAAAGAGCTTCAAAAGAGTCAAGATAAACTTCAGCA 291
Qy 258 ACCCTTATCTGACGCGATGCCAGGCTGCGGTGCAACAGACCGAGTTCCTGGGGAAG 317
Db 292 ACCCTTCTCGCAGCAGATGCCAGGCTCCAGTGCATAAGACTGAGTTTCTGGGAAAG 351
Qy 318 AATGAAGTTGATTTTGTCTCAGACTTTTACATAGGAAGTTTACACCTGGCTCCGCAA 377
Db 352 AATGAAGTTTATATTTTGTCTCAGACCTTACATAGGAAGTCAACCTGGCTCCGCAA 411
Qy 378 TCCGCAAAACAGTTCCTCATCTCCCTCCGGCTCTCTCCGTTGGTGGGAAACAAGT 437
Db 412 TCCAGACAGCAGTTTCTGATCTCCCTCCGCTCTCCGSCAGTGGGATGGAACAAGT 471
Qy 438 AGAAGATGCCACCCCTGTCATAAAATACGATCTTTTATATGCCATCTCCAAGCTGGGCC 497
Db 472 GGAAGATGCCACCCCTGTCATAAACTATGATCTTTATATGCCATCTCCAAGCTGGGCC 531
Qy 498 AGGAGAGATATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 557
Db 532 AGGGGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 591
Qy 558 GTGTGAGAGTGACCAAGAGAAATGAGGAGGAAGGAA 594
Db 592 ATGTRAGATGATCAAGAGAAAGGAGGAAGAGGAA 628

RESULT 7
AAD30156
ID AAD30156 standard; DNA; 2212 BP.
XX
XX
AC AAD30156;

XX 17-MAY-2002 (first entry)
 DT Human MCIP associated DNA #2.
 DE
 XX Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;
 KM heart failure; cardiomyopathy; heart disease; human; gene; ds.
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FH 25..618
 FT CDS /*tag= a
 FT /product= "Human MCIP associated protein #2"
 PN WO200204491-A2.
 XX 17-JAN-2002.
 PD
 PF 06-JUL-2001; 2001MO-US21662.
 XX
 PR 07-JUL-2000; 2000US-216601P.
 PR 13-FEB-2001; 2001US-0782953.
 XX
 PA (TEXA) UNIV TEXAS SYSTEM.
 PA (WILL/) WILLIAMS S R.
 PA (ROTH/) ROTHERMEL B.
 PI Williams SR, Rothermel B;
 XX WPI: 2002-179698/23.
 DR F-PSDB; AAE18915.
 XX
 PT Screening for modulators of muscle calcineurin interacting protein
 PT (MCIP) binding, expression or phosphorylation, useful for treating
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,
 PT calcineurin and a test compound -
 XX
 XX Example 1; Page 159-161; 174pp; English.
 PS
 XX The invention relates to muscle calcineurin interacting proteins (MCIPs)
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical
 CC complex with the catalytic subunit of calcineurin and increased levels
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate
 CC transcription of certain target genes. The invention also relates to
 CC methods for identifying modulators of MCIP binding, expression or
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin
 CC may be used for treating cardiac hypertrophy and heart failure.
 CC Antibodies to MCIP can be used in characterizing the MCIP content of
 CC healthy and diseased tissues and subsequently for determining the
 CC presence or absence of cardiomyopathy or as predictor of heart disease.
 CC Note: This sequence has been described as murine MCIP splice variant in
 CC the specification, however the sequence seems to be a polynucleotide
 CC encoding a MCIP associated protein.
 CC
 XX Sequence 2212 BP; 588 A; 455 C; 529 G; 640 T; 0 other;
 SQ
 XX Query Match 58.3%; Score 349.2; DB 24; Length 2212;
 XX Best Local Similarity 80.9%; Pred. No. 1.8e-102;
 XX Matches 419; Conservative 0; Mismatches 98; Indels 1; Gaps 1;
 OY 78 AGAATGATTTAGGACCTTACCAATTTAGCTCCCTGATTCCTGTGCGAAACG 137
 DB 23 AGATGGAGAGGTGACCTCGACGACCTGCCAGGCCACATGCGCTGCACCTGGACC 82
 OY 138 ATGATGCTTTCAGCGAAAGTACAGACGCGCCAAATTTGATCCTCTTTCAGAACTATG 197
 DB 83 CGCGGTGTTCGTGACGCGCTGTGCGCGCCAAATTTGAGTCCCTCTTTGGAGAGTATG 142
 OY 198 ACAAGACACACCTTCCGATATTTTAAAGCTTCAACGCTGTCGGATTAACCTCAGCA 257
 DB 143 ACAAGACATCACCTTTCAGTATTTTAAAGAGCTTCAACGAGTCAGAATTAACCTCAGCA 202

OY 258 ACCCTTATCTGACCGCATGCGAGCTCGGCTGCAACAACCGATTCTTGGGAAAG 317
 DB 203 ACCCTTCTCCGACGAGATGCGAGCTCCAGCTGCATTAAGATGTTCTGGGAAAG 262
 OY 318 AATGAAGTTGATTTTTCCTCAGACTTTACATAGAAAGTTACACACTGGCTCCGCC-A 376
 DB 263 AATGAAGTTATTTTTCCTCAGACTTTACATAGAAAGTTACACACTGGCTCCGCCA 322
 OY 377 ATCCGACAAACAGTTCCCTCATCTCCCTCCGACCTCTCCTCCGTTGGCTGAAACAAG 436
 DB 323 ATCCGACAAACAGTTCTCATCTCCCTCCGACCTCTCCTCCGAGTGGATGAAACAAG 382
 OY 437 TAGAAGATGCCACCCCGCTATTAATTAATTCATCTTTTATTTGCACTCTCCAACTGGGCG 496
 DB 383 TGAAGATGGACCCCGCTATTAATTAATTCATCTTTTATTTGCACTCTCCAACTGGGCG 442
 OY 497 CAGAGAGAAAGTATGAACCTCATGCGAGCAGACCACTCCAGTGTGTGTCACG 556
 DB 443 CAGGGAAGAAAGTATGAATTCGACGCGAGCTGACACCACTCCAGGTGTGTGTCATG 502
 OY 557 TGTGTGAGAGTGAACCAAGAGATGAGAGAGAGAGAA 594
 DB 503 TATGTGAGAGTGAACCAAGAGAGAGAGAGAGAA 540
 RESULT 8
 ABN96843
 ID ABN96843 standard; DNA; 2173 BP.
 XX
 AC ABN96843;
 XX
 DT 13-AUG-2002 (first entry)
 XX
 DE Gene #3341 used to diagnose liver cancer.
 DE
 KM Gene; liver cancer; ds; hepatocellular carcinoma; hepatocytic;
 KM metastatic liver tumour; cytostatic; expression profile; disease state;
 KM disease progression; drug toxicity; drug efficacy; drug metabolism.
 XX
 OS Homo sapiens.
 XX
 XX WO200229103-A2.
 PN
 XX 11-APR-2002.
 PD
 PF 02-OCT-2001; 2001MO-US30589.
 XX
 PR 02-OCT-2000; 2000US-237054P.
 XX
 PA (GENE-) GENE LOGIC INC.
 XX
 PI Horne D, Alvares C, Pires-da-Silva S, Vockley JG;
 XX WPI: 2002-426119/45.
 DR
 XX
 PT Diagnosing and detecting the progression of liver cancer.
 PT hepatocellular carcinoma or metastatic liver tumor in a patient.
 PT involves detecting the level of expression of two or more genes in a
 PT liver tissue sample -
 XX
 PS Claim 1; SEQ ID NO 3341; 298bp; English.
 XX
 XX The invention relates to a novel method for diagnosing and detecting the
 CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
 CC tumour in a patient, and differentiating metastatic liver cancer from
 CC hepatocellular carcinoma in a patient, involving detecting the level of
 CC expression of two or more genes represented in ABN93503-ABN97455 in a
 CC tissue sample. The method of the invention has hepatocytic, and
 CC cytostatic activity. The method is useful for diagnosing and detecting
 CC the progression of liver cancer, hepatocellular carcinoma and metastatic
 CC liver carcinoma in a patient. The method is useful for identifying
 CC expression profiles which serve as useful diagnostic markers as well as

CC markers that can be used to monitor disease states, disease progression,
CC drug toxicity, drug efficacy and drug metabolism.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2173 BP; 606 A; 431 C; 503 G; 633 T; 0 other;

Query Match 57.1%; Score 341.8; DB 24; Length 2173;
Best Local Similarity 88.6%; Pred. No. 4.4e-100;
Matches 381; Conservative 0; Mismatches 47; Indels 1; Gaps 1;

QY 167 GCCAAATTTGAATCCCTCTTCAGACATATGACAAGACACACCTTCCAGTATTTTAAG 226
DB 58 CCCAAATTTGAGTCCCTCTTTAGGACGTATGACAAGGACATCACCTTTTCAGTATTTTAAG 117
QY 227 AGCTTCAAACGTGTCGGGATAACTTCAGCAACCCCTTATCTGCAGCCGATGCCAGGCTG 286
DB 118 AGCTTCAAACGTGTCAGATAAATCTTCAGCAACCCCTTCTCCGACGAGATGCCAGGCTC 177
QY 287 CGGCTGCACAAGACCGAGTTCCTGGGGAAGAAATGAAGTTGTATTTTGTCTCAGACTTTA 346
DB 178 CAGCTGCATAAGACTGAGTTTCTGGGAAGAAATGAAGTTATATTTTGTCTCAGACTTTA 237
QY 347 CACATAGGAGTTTCAACCTGGCTCCGCC-AATCCCGACAAACAGTTCTCATCTCCCT 405
DB 238 CACATAGGAGTTTCAACCTGGCTCCGCCAAATCCAGCAAGCAGTTTCTGATCTCCCT 297
QY 406 CCGGCTCTCTCCGTTGGTGGTGAACAAGTAGAAGTACACCCCGTCATATAATTAC 465
DB 298 CCGGCTCTCCGAGTGGATGGAAACAGTGGAGTCCGACCCAGTATGAATTTGACAGCG 357
QY 466 GATCTTTTATATGCGATCTCAAGCTGGGCGAGGAGAGTATGAATGTCATGCGCG 525
DB 358 GATCTTTTATATGCGATCTCAAGCTGGGCGAGGAGTGGAGTACCAAGAGATGAGGAG 585
QY 526 ACAGACACCACTCCAGTGGTGGTCCAGTGTGTGAGAGTACCAAGAGATGAGGAG 585
DB 418 ACTGACACCACTCCAGTGGTGGTCCATGTATGTGAGAGTATCAAGAGAGAGGAA 477
QY 586 GAAGAGGAA 594
DB 478 GAAGAGGAA 486

RESULT 9
ID AAX01282 standard; cDNA to mRNA; 2174 BP.
AC AAX01282;
XX 09-APR-1999 (first entry)
XX Human DSCR1 coding sequence.
DE DSCR1; Down's Syndrome critical region 1; human; pathogenic abnormality;
KW Central Nervous System development; mental retardation; heart defect; ds.
XX Homo sapiens.
FH Key Location/Qualifiers
FT CDS 49..564
FT /*tag= a
FT polyA_signal 1541..1546
FT /*tag= b
FT polyA_signal 2132..2137
FT /*tag= c
XX US5869318-A.
XX 09-FEB-1999.
XX 07-JUN-1996; 96US-0665040.

XX 07-JUN-1995; 95ES-0001140.
PR (PALL/) PALLEJA X E.
XX Puentes JJ, Palleja XE, Pritchard M;
PI WPI: 1999-152781/13.
DR P-PSDB; AAW73898.
XX DNA encoding foetal brain proteins - believed to be associated with
PT Down's syndrome
PS Claim 1; Column 15-18; 19pp; English.
XX This sequence is the Down's Syndrome critical region 1 (DSCR1) gene
CC of the invention. The DSCR1 gene was found to be located in the
CC q22.1-22.2 region of human chromosome 21. An increase in the transient
CC expression of DSCR1 mRNA in the brains of young rats, compared to
CC expression levels in the brains of adult rats, suggests an important role
CC for DSCR1 during the development of the Central Nervous System (CNS), and
CC that over expression of DSCR1 may be involved in pathogenic abnormalities
CC of mental retardation and/or heart defects as found in Down's syndrome
CC patients.
XX Sequence 2174 BP; 605 A; 436 C; 502 G; 628 T; 3 other;

Query Match 56.8%; Score 340.2; DB 20; Length 2174;
Best Local Similarity 88.6%; Pred. No. 1.5e-99;
Matches 380; Conservative 0; Mismatches 48; Indels 1; Gaps 1;
QY 167 GCCAAATTTGAATCCCTCTTCAGACATATGACAAGGACACACCTTCCAGTATTTTAAG 226
DB 58 GCCAAATTTGAGTCCCTCTTTAGGACGTATGACAAGGACATCACCTTTTCAGTATTTTAAG 117
QY 227 AGCTTCAAACGTGTCGGGATAACTTCAGCAACCCCTTATCTGCAGCCGATGCCAGGCTG 286
DB 118 AGCTTCAAACGTGTCAGATAAATCTTCAGCAACCCCTTCTCCGACGAGATGCCAGGCTC 177
QY 287 CGGCTGCACAAGACCGAGTTCCTGGGGAAGAAATGAAGTTGTATTTTGTCTCAGACTTTA 346
DB 178 CAGCTGCATAAGACTGAGTTTCTGGGAAGAAATGAAGTTATATTTTGTCTCAGACTTTA 237
QY 347 CACATAGGAGTTTCAACCTGGCTCCGCC-AATCCCGACAAACAGTTCTCATCTCCCT 405
DB 238 CACATAGGAGTTTCAACCTGGCTCCGCCAAATCCAGCAAGCAGTTTCTGATCTCCCT 297
QY 406 CCGGCTCTCTCCGTTGGTGGTGAACAAGTAGAAGTACACCCCGTCATATAATTAC 465
DB 298 CCGGCTCTCCGAGTGGATGGAAACAGTGGAGTCCGACCCAGTATGAATTTGACAGCG 357
QY 466 GATCTTTTATATGCGATCTCAAGCTGGGCGAGGAGAGTATGAATGTCATGCGCG 525
DB 358 GATCTTTTATATGCGATCTCAAGCTGGGCGAGGAGTGGAGTACCAAGAGATGAGGAG 417
QY 526 ACAGACACCACTCCAGTGGTGGTCCAGTGTGTGAGAGTACCAAGAGATGAGGAG 585
DB 418 ACTGACACCACTCCAGTGGTGGTCCATGTATGTGAGAGTATCAAGAGAGAGGAA 477
QY 586 GAAGAGGAA 594
DB 478 GAAGAGGAA 486

RESULT 10
AAC01774
ID AAC01774 standard; cDNA; 442 BP.
XX AAC01774;
XX 06-OCT-2000 (first entry)
XX Human secreted protein 5' EST, SEQ ID NO: 1772.

XX	Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX	gene therapy; chromosome mapping; ss.
XX	
XX	Homo sapiens.
XX	
XX	EP1033401-A2.
XX	
XX	06-SEP-2000.
XX	
XX	21-FEB-2000; 2000EP-0200610.
XX	
XX	26-FEB-1999; 99US-0122487.
XX	
XX	(GEST) GENSET.
XX	
XX	Dumas Mline Edwards J, Duclert A, Giordano J;
XX	
XX	WPI; 2000-500381/45.
XX	P-PSDB; AAG01/68.
XX	
XX	New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX	obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
XX	diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX	
XX	Claim 1, SEQ ID 1772; 71bp + CD-ROM; English.
XX	
XX	The present sequence is one of a large number of 5' ESTs derived from
XX	mRNAs encoding secreted proteins. An ORF has been identified within the
XX	sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
XX	derived from 30 different tissues. EST sequences usually correspond
XX	mainly to the 3' untranslated region (UTR) of the mRNA because they are
XX	often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
XX	well suited for isolating cDNA sequences derived from the 5' ends of
XX	mRNAs and even in those cases where longer cDNA sequences have been
XX	obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
XX	mRNAs with intact 5' ends and can therefore be used to obtain full length
XX	cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
XX	gene therapy and chromosome mapping procedures. They are used to obtain
XX	upstream regulatory sequences and to design expression and secretion
XX	vectors.
XX	
XX	Sequence 442 BP; 127 A; 111 C; 93 G; 111 T; 0 other;
XX	
XX	Query Match 44.1%; Score 264.4; DB 21; Length 442;
XX	Best Local Similarity 83.4%; Pred. No. 2.5e-75;
XX	Matches 312; Conservative 0; Mismatches 61; Indels 1; Gaps 1

[illegible]

Db 210 CTGTGATGTTTCCACTCTGGTTGCTGTGGTGGATGTCGAGGTCTTTACCAATCAGGA 269
Qy 160 GACGAGGCCCCAAATTTGTAATCCCTCTTCAGAAATATATGACAAAGGACACCACTTCCAGTA 219
Db 270 GGTAAAGGAAAAATTTGGGGACTGTTTCGGACTTATGATGACTGTGTGAGGTTCCAGCT 329
Qy 220 TTTTAAAGAGCTTCAACGTCGCGGATAACTTCAGCAACCCCTTATCTGAGCCGATGC 279
Db 330 ATTTAAGAGTTTTCAGACGTCGCGTATAAACTTCAGCAATCTTAAATCTGCAGCCGAGC 389
Qy 280 CAGGTCGGCTGCACAAAGACCGAGTCTCTGGGCAAGCAATGAAGTTGTTATTTGCTCA 339
Db 390 TAGGATAGAGCTTCATGAACCCCAATTCAGAGGGGAAAAATTAAGCTCTACTTTGCACA 449
Qy 340 GACT-----TTACACATAGGAAGTTTCACACTGGCTCC-GCCAAATCCCGACAA 386
Db 450 GGTTCAGACTCCAGACAGACAGATGAGACAAACTGCACCTTGGCTCCACCCGAGCTGCCAA 509
Qy 387 ACAGTCTCTATCTCCCTCCGGCTCTCTCTCCGTTGGCTGGAAACAAGTAGAAGATGC 446
Db 510 ACAGTCTCTATCTCGCCCTCTCTCTCCACCTGTTAGCTGGCAGCCCATCAACGATGC 569
Qy 447 CACCCCGCTCATAAATACGATCTTTTATATGCCATCTCCAAGCTGGGCGCAGGAGAA 506
Db 570 CAGCCAGTCTCAACTATGACCTCTCTATGCTGTGGCCAACTAGGACCCAGAGAGAA 629
Qy 507 GTATGAAGTGCATGACGAGCAGACACACTCCAGTGTGGTGGTCCAGCTGTGTGAGAG 566
Db 630 GTATGAGCTCATGACGAGGACTGAGTCCACCCCAAGTGTGCTGTGACGCTGTGCAGAC 689
Qy 567 TGACCAAGAGAAATGAGGAGGA 587
Db 690 TGACATAGAGAAAGAGGA 710

RESULT 12
ABL61768
ID ABL61768 standard; DNA; 3184 BP.
AC ABL61768;
XX
DT 15-MAY-2002 (first entry)
XX
DE Colon adenocarcinoma related gene sequence SEQ ID NO:105.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
PD
XX
XX 13-DEC-2001.
XX
XX 30-MAY-2001; 2001WO-US10838.
XX
PR 05-JUN-2000; 2000US-209473P.
PR 05-JUN-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-231133P.
PR 18-SEP-2000; 2000US-233617P.
PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.

PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 29-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.
PA (AVAL-) AVALON PHARM.
XX
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
XX WPI; 2002-188264/24.
XX
PT Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set -
XX
XX Claim 1; SEQ ID 105; 44pp; English.
XX
CC The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilm's tumour.
XX
SQ Sequence 3184 BP; 921 A; 681 C; 657 G; 925 T; 0 other;
Query Match 33.8%; Score 202.2; DB 24; Length 3184;
Best Local Similarity 65.9%; Pred No. 1.1e-54;
Matches 330; Conservative 0; Mismatches 158; Indels 13; Gaps 2;
Qy 100 CTACAATTTAGCTCCCTGATTTGCTGTGGCAACGATGATCTTTCAGCGAAAGTGA 159
Db 210 CTGTGATGTTTCCACTCTGGTTGCTGTGGTGGATGTCGAGGTCTTTACCAATCAGGA 269
Qy 160 GACGAGGCCCCAAATTTGTAATCCCTCTTCAGAAATATGACAAAGGACACCACTTCCAGTA 219
Db 270 GGTAAAGGAAAAATTTGGGGACTGTTTCGGACTTATGATGACTGTGTGAGGTTCCAGCT 329

Db 390 TAGATAGAGTTTCATGAACCCCAATTCAGAGGGGAAAAATTAAGCTCTACTTTGCACA 449
 QY 340 GACT-----TTACACATAGGAAGTTTCACCTGGCTCC-GCCAAATCCCGACAA 386
 Db 450 GGTTCAGACTCCAGAGACAGATGGAGACAAACTGCACCTTGGCTCCACCCAGCTGCCAA 509
 QY 387 ACAGTTCTCATCTCCCTCGGCTCTCTCCGGTTGGCTGGAACAGTAGAAGATGC 446
 Db 510 ACAGTTTCTCATCTCGGCTCCCTCCCTCCACCTGTTAGCTGCGAGCCCACTCAACGATGC 569
 QY 447 CACCCCGCTCATAAATPACCATCTTTTATATGTCATCTCCAAAGCTGGGCGCAGGAGAA 506
 Db 570 CAGCCAGTCTCACTATGACTCTCTCTATGCTGTGGCCAACTAGGACACAGAGAGAA 629
 QY 507 GTATGAATCGATCGCGGACAGACACCACTCCCAAGTGTGGTCCAGTGTGTGAGAG 566
 Db 630 GTATGAGTCCATCGAGGACTGAGTCCACCCCAAGTGTGCTGTCAGCTGTGCGACAG 689
 QY 567 TGACCAAGAGAATGAGGAGGA 587
 Db 690 TGACATAGAGGAAGAGAGGA 710

RESULT 14

AAF25338

ID AAF25338 standard; cDNA; 934 BP.

XX AAF25338;

AC AAF25338;

DT 30-APR-2001 (first entry)

XX Nucleotide sequence of a human detoxification protein.

DE Human; detoxification protein; DETX; cancer; leukaemia; melanoma;
 KW adenocarcinoma; autoimmune disorder; inflammatory disorder;
 KW rheumatoid arthritis; asthma; atherosclerosis; autoimmune thyroiditis;
 KW psoriasis; ulcerative colitis; infection; cell proliferative disorder;
 KW actinic keratosis; arteriosclerosis; cirrhosis; hepatitis; ss.

XX Homo sapiens.

XX Key

FH Location/Qualifiers

FT 20..724

FT /*tag= a

FT /product= "detoxification protein"

FT 20..100

FT /*tag= b

FT sig_peptide

FT WO200104305-A2.

XX 18-JAN-2001.

XX 06-JUL-2000; 2000WO-US18509.

XX 07-JUL-1999; 99US-0142678.

XX (INCYTE) INCYTE GENOMICS INC.

XX Tang YT, Yue H;

XX WPI; 2001-147193/15.

XX P-PSDB; AAB31788.

XX New human detoxification protein and polynucleotide, useful for

XX diagnosis, prevention and treatment of autoimmune/inflammatory

XX disorders and cell proliferative disorders including cancer

XX Claim 5; Page 79; 79pp; English.

XX The present sequence encodes a human detoxification polypeptide (DETX).

XX DETX and its (ant)agonists are useful for preventing or treating

XX disorders associated with decreased or increased expression or activity

XX of DETX. DETX polypeptides are useful for screening compounds that

CC specifically binds to DETX and for identifying (ant)agonists.
 CC Diseases prevented, treated and diagnosed include cancers (e.g.
 CC leukaemia, melanoma, adenocarcinoma, cancers of the brain, bone marrow,
 CC breast, kidney, liver, pancreas, prostate and uterus),
 CC autoimmune/inflammatory disorders (e.g. rheumatoid arthritis, asthma,
 CC atherosclerosis, autoimmune thyroiditis, psoriasis and ulcerative
 CC colitis), bacterial, fungal, parasitic infections and cell
 CC proliferative disorders (e.g. actinic keratosis, arteriosclerosis,
 CC cirrhosis and hepatitis). Anti-DETX antibodies may be used as
 CC antagonists, as a targeting or delivery mechanism for bringing
 CC pharmaceutical agents into contact with cells or tissues expressing
 CC DETX and for diagnosis of DETX-related disorders.

XX
 SQ Sequence 934 BP; 247 A; 244 C; 206 G; 237 T; 0 other;

Query Match 33.0%; Score 197.6; DB 22; Length 934;

Best Local Similarity 64.6%; Pred. No. 1.8e-53;

Matches 332; Conservative 0; Mismatches 169; Indels 13; Gaps 2;

QY 87 TTAGGGACTTTAGCTACAAATTTTAGCTCCCTGATTGCTTGTGGCAACGATGATGCT 146
 Db 138 TCACTGACTTCAATGACCTCCCAACTCGTTGTTTGGCGCAATGTTCCACGATCAGTGT 197
 QY 147 TCAGCGAAAGTGAGACCGAGCGCCAAATTTGAATCCCTCTTCAGAAATATGCAAGGACA 206
 Db 198 TTGAAGGAGAGAGAGCAAGGAAAAATTTGAGGGACTGTTTCGGACTTATGATGACTGTG 257
 QY 207 CCACCTTCCAGTATTTTAAGAGCTTCAACGCTGTCGGATAACTTCAGCAACCCCTTAT 266
 Db 258 TGACGTTCAGCTATTTAAGAGTTCAGACGTTCCGTATATAAACTTCAGCAATCTCTAAAT 317
 QY 267 CTGACGCCGATGCCAGGTGCGGCTGCACAGACCGAGTTCCTGGGAGGAAATGAAGT 326
 Db 318 CTGACGCCGAGCTAGGATAGAGCTTCATGAACCCCAATTTCAGAGGGAAAAATTAAGC 377
 QY 327 TGTATTTTGTCTCAGACT-----TTACACATAGGAAGTTTCACACTGGGTCC-G 373
 Db 378 TCTACTTTTGCACAGGTTTCAGACTCCAGAGACAGATGGAGACAAACTGCACTTGGCTCCAC 437
 QY 374 CCAATCCCGACAAACAGTTTCCTCATCTCCCTCCGGCTCTCTCCCTGGCTGGAAC 433
 Db 438 CCCAGCTTCCCAACAGTTTTCATCTCGCCCCCTTCCTCCCACTGTGTGGCTGGCAGC 497
 QY 434 AAGTAGAAGATGCCACCCCGTCATAAATTCAGATCTTTTATATGCCATCTCCAAAGCTGG 493
 Db 498 CCATCAACGATGCCACCGCCAGTCCCTCACTATGACTCTCTATGCTGTGGCCAACTAG 557
 QY 494 GGCAGGAGAGAGTATGAACCTGCATGCGCAGCAGACACCACTCCCACTGTGGTGGTCC 553
 Db 558 GACCAGGAGAGAGTATGAGCTCCATGCGAGGAGTGAAGTCCACCCCAAGTGTGCTGTGC 617
 QY 554 ACGTGTGTGAGAGTGACCAAGAGATGAGGAGGA 587
 Db 618 ACGTGTGCACAGTGACATAGAGGAAGAAGAGGA 651

RESULT 15

ABA91457

ID ABA91457 standard; cDNA; 3159 BP.

XX ABA91457;

AC ABA91457;

XX 18-APR-2002 (first entry)

DT Human Down syndrome critical region 1-like 1 protein cDNA.

DE Down syndrome critical region 1-like 1; DSCR1L1 alpha; human;

XX Down syndrome; Alzheimer's disease; dementia; transgenic;

KW neuroprotective; neurotropic; anticonvulsant; diagnosis;

KW gene therapy; gene; ss.

XX Homo sapiens.

OS Homo sapiens.

XX

```

FH Key Location/Qualifiers
FT CDS 185..952
FT /tag=a
FT /product="DSCR1L1"
FT /trans_except="(pos:215..217, aa:Xaa)
FT /note="Xaa = unknown"
FN MO200204513-A2.
XX
XX 17-JAN-2002.
PD
PD 11-JUL-2001; 2001WO-US21982.
PF
PF 11-JUL-2000; 2000US-0614474.
PR
PR (INCY-) INCYTE GENOMICS INC.
PA
PA Loring JF, Tingley DW, Edwards CM, Streeter DG;
PI MPI; 2002-164633/21.
XX P-PSDB; AAM50760.
DR
DR Novel Down syndrome critical region 1-like 1 protein and nucleic acid
PT encoding the protein useful for diagnosis and treatment of Alzheimer's
PT disease, Down syndrome and other forms of dementia
XX
XX PS Claim 2; Page 46-47; 54pp; English.
XX
XX CC The present sequence is that of cDNA clone Incyte ID No: 247500.5
CC encoding novel Down syndrome critical region 1-like 1 protein
CC (DSCR1L1 alpha, see AAM50760). Northern analysis indicated
CC expression of DSCR1L1 alpha in various libraries, with the highest
CC abundance in tissues from the nervous system, including tissues
CC associated with schizophrenia, Huntington's disease, epilepsy and
CC amyotrophic lateral sclerosis. An absence of DSCR1L1 expression
CC was observed in 7 of 8 libraries from subjects with Alzheimer's
CC disease. A claimed method for detecting differential expression of
CC a nucleic acid encoding DSCR1L1 alpha can be used to diagnose Down
CC syndrome, Alzheimer's disease and other forms of dementia. A probe
CC from such a nucleic acid is useful for identifying naturally
CC occurring molecules encoding DSCR1L1 alpha, allelic variants or
CC other molecules. The nucleic acids are useful for producing
CC transgenic cell lines or organisms which model human disorders.
CC They may also be used in gene therapy, and to screen for ligands
CC which specifically bind the nucleic acid molecule, such as
CC a DNA or RNA molecule, peptide nucleic acid, artificial chromosome,
CC peptide, transcription factor, repressor or regulatory molecule,
CC for use as a therapeutic.
XX
XX SQ Sequence 3159 BP; 926 A; 657 C; 657 G; 915 T; 4 other;
XX
XX Query Match 33.0%; Score 197.6; DB 24; Length 3159;
XX Best Local Similarity 64.8%; Pred. No. 3.4e-53;
XX Matches 332; Conservative 0; Mismatches 169; Indels 13; Gaps 2;
XX
OY 87 TTAGGGAATTACTTCAATTTAGTCTCCCTGATTTGCTGTGGCAACGATGATGCTT 146
DB 366 TCACCTAATTCAATGACCTCCCACTCGTTGTTGGTGCAATGTTCAACGATGATGCT 425
OY 147 TCAGCGAAGTGAAGACGAGGCGCAATTTGAATCCCTCTTCAAGAACATATGACAGGACA 206
DB 426 TTGAGGAGAAAGAGAGCAAAATTTGAGGAGCTGTTCCGACTATGATGACTGTG 485
OY 207 CCACCTTCCAGTATTTTAAGAGCTTCAAGCTGTCGCGAATAAATTGACGACCCCTTAT 266
DB 486 TGAAGTTCAGCTATTTAAGAGTTTCAAGCTGTCGCGAATAAATTGACGACCCCTTAT 545
OY 267 CTGAGGCGATGCGAGGCTGCGGCTGCAAGACCGAGTTCTGCGGAGAAATGAGT 326
DB 546 CTGAGGCGCGAGGCTGAGATGAGCTTCAATGAAACCAATTCAGAGGAGAAATTAAGC 605
OY 327 TGTATTTTGTCTCAGACT-----TTACATAGGAAGTTCAACCTGGCTCC-G 373

```

```

DB 606 TCTACTTTGCACAGGTTCAAGCTCCAGAGACAGATGAGAGCAAACTGACCTGGCTCCAC 665
OY 374 CCAATCCCGACAAACAGTTCTCATCTCCCTCCGACCTCTCTCCGTTGGCTGGAAC 433
DB 666 CCGAGCTGCGCAAAACAGTTTCTCATCTCGCCCTCTCTCCCACTGTTGGCTGGGAGC 725
OY 434 AAGTAGAAGATGCCACCCCGCTCATTAATTAAGATCTTTATATGCGATGCCAAGCTGG 493
DB 726 CCATCAAGATGCCACCGCACTCTCTCACTATGACCTCTCTATGCTGGCCAAACTAG 785
OY 494 GGCCAGAGAGAAATGATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 553
DB 786 GACCAAGAGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 845
OY 554 ACGTGTGTGAGAGTGAACCAAGAGATGAGAGGA 587
DB 846 ACGTGTGTGAGAGTGAACCAAGAGATGAGAGGA 879

```

Search completed: December 14, 2002, 15:18:56
Job time : 103.824 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 14, 2002, 18:11:26 ; Search time 139.234 Seconds
(without alignments)
3202.482 Million cell updates/sec

Title: US-09-782-953-7

Perfect score: 1043

Sequence: 1 MDRFRSYNPSLIACVAND.....RPKPKIQTTRRPTPIHLS 198

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09782953/runat_11122002_114428_17330/app_query.fasta_1.1173
-DB=N Geneseq 101002 -OPMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09782953 @CGN_1_1220 @runat_11122002_114428_17330 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq 101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1043	100.0	597	24	AAD30153	Mouse MCIP associa
2	970.5	93.0	2331	24	AAD30155	Human MCIP associa
3	957.5	91.8	2348	24	AAS94805	Human DNA sequence
4	920.5	88.3	2597	24	AAD30152	Human MCIP associa
5	875	83.9	2212	24	AAD30156	Human MCIP associa
6	868	83.2	599	24	AAD30151	Human MCIP associa
7	847.5	81.3	2173	24	AEN96843	Gene #3341 used to
8	844.5	81.0	2174	20	AAF101282	Human DSCR1 coding
9	830	79.6	2358	24	AAF18328	Lung cancer associ
10	632.5	60.6	594	24	AAD30154	Mouse MCIP associa
11	623	59.7	934	22	AAF25338	Nucleotide sequenc
12	623	59.7	3159	24	ABA91457	Human Down syndrom
13	615.5	59.0	3184	24	AAD30157	Human MCIP associa
14	615.5	59.0	3184	24	ABL61768	Colon adenocarcino
15	615.5	59.0	3184	24	ABL65187	Lung cancer relate
16	592.5	56.8	828	24	AAD30158	Human MCIP associa
17	591.5	56.7	720	24	AAD30159	Human secreted pro
18	537	51.5	442	21	AAC01774	Rat Down syndrome
19	536.5	51.4	615	24	ABA91463	Rat Down syndrome
20	528.5	50.7	1021	23	ABA91458	Drosophila melanog
21	372	35.7	1820	23	ABL29301	Human foetal liver
22	372	35.7	12550	23	ABL29300	Human brain expres
23	285.5	27.4	412	22	ABA56172	Human foetal liver
24	285.5	27.4	412	22	AAK04364	Human brain expres
25	285.5	27.4	412	22	AAI14449	Probe #4382 for ge
26	285.5	27.4	412	22	AAI35821	Probe #4507 used t
27	285.5	27.4	412	22	AAI04272	Probe #4263 used t
28	285.5	27.4	412	22	ABO4419	Human genome-deriv
29	285.5	27.4	446	22	ABA43406	Human breast cell
30	285.5	27.4	446	22	ABA33852	Human foetal liver
31	285.5	27.4	446	22	ABA23598	Probe #2064 for ge
32	285.5	27.4	446	22	AAK02113	Human brain expres
33	285.5	27.4	446	22	AAK27563	Human bone marrow
34	285.5	27.4	446	22	AAI12148	Probe #2081 for ge
35	285.5	27.4	446	22	AAI33493	Probe #2179 used t
36	285.5	27.4	446	22	AAI02061	Probe #2052 used t
37	285.5	27.4	446	24	ABO52041	Human genome-deriv
38	272	26.1	640	24	ABO59719	Human colon cancer
39	266.5	25.6	486	22	ABA42132	Human breast cell
40	266.5	25.6	486	22	ABA25254	Human foetal liver
41	266.5	25.6	486	22	ABA22343	Probe #809 for gen
42	266.5	25.6	486	22	AAK00816	Human brain expres
43	266.5	25.6	486	22	AAK26270	Human bone marrow
44	266.5	25.6	486	22	AAI10903	Probe #836 for gen
45	266.5	25.6	486	22	AAI32163	Probe #849 used to

ALIGNMENTS

RESULT 1

AAD30153

ID AAD30153 standard; DNA; 597 BP.

XX AAD30153;

XX 17-MAY-2002 (first entry)

DT Mouse MCIP associated DNA #3.

DE Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;

KW heart failure; cardiomyopathy; heart disease; mouse; gene; ds.

XX Mus musculus.

XX Key Location/Qualifiers

FT CDS 1..597

FT /*tag= a

FT /product= "Mouse MCIP associated protein #3"

XX WO200204491-A2.

XX 17-JAN-2002.

XX 06-JUL-2001; 2001WO-US21662.

XX 07-JUL-2000; 2000US-216601P.

XX 13-FEB-2001; 2001US-0782953.

XX (TEXA) UNIV TEXAS SYSTEM.

XX (WILL/) WILLIAMS S R.

XX (ROTH/) ROTHERMEL B.

XX Williams SR, Rothermel B;

XX WPI, 2002-179698/23.

XX P-Psdb; AAE18912.

XX Screening for modulators of muscle calcineurin interacting protein

XX (MCIP) binding, expression or phosphorylation, useful for treating

XX cardiac hypertrophy or heart failure, comprises mixing MCIP,

XX calcineurin and a test compound -

XX Disclosure: Page 150-151; 174pp; English.

XX The invention relates to muscle calcineurin interacting proteins (MCIPs)

XX and nucleic acid molecules encoding such proteins. MCIPs form a physical

XX complex with the catalytic subunit of calcineurin and increased levels

XX of MCIPs correspond to a reduced ability of calcineurin to stimulate

XX transcription of certain target genes. The invention also relates to

XX methods for identifying modulators of MCIP binding, expression or

XX phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin

XX may be used for treating cardiac hypertrophy and heart failure.

XX Antibodies to MCIP can be used in characterizing the MCIP content of

XX healthy and diseased tissues and subsequently for determining the

XX presence or absence of cardiomyopathy or as predictor of heart disease.

XX The present sequence is mouse MCIP associated DNA.

XX Note: This sequence has been described as human MCIP3 encoding DNA

XX in the specification, however the sequence seems to be a polynucleotide

XX encoding a MCIP associated protein.

XX SQ Sequence 597 BP; 170 A; 156 C; 142 G; 129 T; 0 other;

XX Alignment Scores:

XX Pred. No.: 4.5e-109 Length: 597

XX Score: 1043.00 Matches: 198

XX Percent Similarity: 100.00% Conservaive: 0

XX Best Local Similarity: 100.00% Mismatches: 0

XX Query Match: 100.00% Indels: 0

XX DB: 24 Gaps: 0

XX US-09-782-953-7 (1-198) x AAD30153 (1-597)

XX 1 MetAspPheArgAspPheSerTyrAsnPheserSerleuIleAcyValAlaAsnAsp 20

XX 1 ATGATTTTGGGACTTACTTCAATTTTAACTCCCTGATTCCTGTCGCAAAAGAT 60

XX 21 AspValPheSerGluSerGluThrArgAlaLysPheGluSerleuPheArgTTrPAsp 40

XX 61 GATGCTTCAGCCAAAGTGAGACCGAGCCAAATTTGAAATCCCTCTTCAGAAATATGAC 120

XX 41 LysAspPheThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheserAsn 60

XX 121 AAGGACCCACCTCCAGTATTTTAAAGCTTCAACGTCTCCGATTAACCTTCAGCAAC 180

XX 61 ProleuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80

XX 181 CCTTATCTCCAGCCGATGCGAGCTCGGCTGCAAAAGACCGAGTCTCGGGAAAGAA 240

XX 81 MetLysLeuTyrPheAlaGlnThrLeuHisLysGlySerSerHisLeuAlaProProAn 100

XX 241 ATGAAGTTGATTTTGTCTCAGACTTACACATAGAAATTCAACCTGCTCCGCCAAT 300

XX 101 ProAspLysGlnPheLeuIleSerProAlaSerProValGlyTyrLysGlnVal 120

XX 301 CCCGACAAACAGTTCCTCACTCCCTCCGCTCTCTCTCCCTGGCTGGAAACAGTA 360

XX 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro 140

XX 361 GAGATGCCACCCCGTCATATAATTAAGATCTTTATATCCCATCTCCAACTGGGGCCA 420

XX 141 GlyGluLysTyrGluLeuHisAlaIleThrAspProThrProSerValValIleHisVal 160

XX 421 GGAGAGAGATGATGACATGACGACGACGACCCCACTCCAGTGTGTTCACGTG 480

XX 161 CysGluSerAspGlnGluAsnGluGluGluGluGluGluGluGluGluGluGluGluGlu 180

XX 481 TGTGAGACTGACCAAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540

XX 181 LysProLysIleIleGlnThrArgArgProGluTyrThrProIleHisLeuSer 198

XX 541 AAGCCAAATCATCCAGACACGAGACCGAGTACACCGATCCACTTAGC 594

XX RESULT 2

XX AAD30155

XX 10 AAD30155 standard; DNA; 2331 BP.

XX AAD30155;

XX 17-MAY-2002 (first entry)

XX Human MCIP associated DNA #1.

XX Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;

XX heart failure; cardiomyopathy; heart disease; human; gene; ds.

XX Homo sapiens.

XX Key Location/Qualifiers

XX FT 144..737

XX FT /*tag= a

XX FT /product= "Human MCIP associated protein #1"

XX WO200204491-A2.

XX 17-JAN-2002.

XX 06-JUL-2001; 2001WO-US21662.

XX 07-JUL-2000; 2000US-216601P.

XX 13-FEB-2001; 2001US-0782953.

XX (TEXA) UNIV TEXAS SYSTEM.

XX (WILL/) WILLIAMS S R.

XX (ROTH/) ROTHERMEL B.

XX Williams SR, Rothermel B;

XX WPI, 2002-179698/23.

XX P-Psdb; AAE18914.

XX Screening for modulators of muscle calcineurin interacting protein

XX (MCIP) binding, expression or phosphorylation, useful for treating

XX cardiac hypertrophy or heart failure, comprises mixing MCIP,

XX calcineurin and a test compound -

XX Claim 72; Page 155-157; 174pp; English.

XX The invention relates to muscle calcineurin interacting proteins (MCIPs)

XX and nucleic acid molecules encoding such proteins. MCIPs form a physical

XX complex with the catalytic subunit of calcineurin and increased levels

XX of MCIPs correspond to a reduced ability of calcineurin to stimulate

XX transcription of certain target genes. The invention also relates to

XX methods for identifying modulators of MCIP binding, expression or

XX	Sequence	599 BP; 164 A; 150 C; 149 G; 136 T; 0 other:
SQ		
Alignment Scores:		
Pred. No.:	3,4e-89	Length: 599
Score:	868.00	Matches: 170
Percent Similarity:	98.84%	Conservative: 0
Best Local Similarity:	98.84%	Mismatch: 2
Query Match:	83.22%	Indels: 1
DB:	24	Gaps: 0
US-09-702-953-7 (1-198) x AAD30151 (1-599)		
OY	2 AAPPheArGAsPheSeSerTyArAnPheSeSerLeuileAlaCyAValAlaAnAspAsp 21	
DB	83 GATTTTAGGACTTAAAGTACAAATTTTAGCTCCGTGATTGCTTGTGTGGCAACATGAT 142	
OY	22 ValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThrTyArgLys 41	
DB	143 GTCCTCAGCGAAAGTGAGAGCCAGGGCCAAATTGGAATCCCTTCAGAACATATACAAAG 202	
OY	42 AAPPThrThrPheGluThrPheLysSerPheLysArgValArgLLeAsnPheSerAsnPro 61	
DB	203 GACACCACCTTCAGATATTTTAAGACCTTCAAAACGTGTCGGATTAACCTTCACGACACC 262	
OY	62 LeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGluMet 81	
DB	263 TTATCTGCAGCCGATCCAGGCTGCGGCTGCACCAAGACCGAGTTCTCGGGGAAGAAATG 322	
OY	82 LysLeuThrPheAlaGlnThrLeuHisLLeGlySerSerHisLeuAlaProProAsnPro 101	
DB	323 AAGTTGATTTTGTCTAGACTTACACATAGGAAGTTTCACACTGCTCCGCCAAT-CCC 381	
OY	102 AAPPLeGlnPheLeuLeuLeuSerProProAlaSerProProValGlyTTrpLysGlnValGlu 121	
DB	382 GACAAACATTTCTCTCATCTCCCTCCGGCTCTCTCCCTGGCTGGCAACAGTAGAA 441	
OY	122 AAPPAlaThrProValLLeAsnTyArgPheLeuLeuTyArgLLeSerLysLeuGlyProGly 141	
DB	442 GATGCCACCCCCCTCATTAATTAATTCATCTTTATATATGATCTCCAAAGCTGGGGCCAGGA 501	
OY	142 GAluLysTyGluLeuHisLAlaAlaThrAspProThrProSerValValHisValCys 161	
DB	502 GAGAACTATGAACTGCATGCATGCAGGACGACAGACACACCTCCCAAGTGTGTGTCCACGTGT 561	
OY	162 GluSerAspGlnGluAsnGluGluGluGluGluGlu 173	
DB	562 GAGAGTGACCAAGAGATGAGGAGGAAGGAAAGAG 597	
RESULT 7		
ABN96843		
ID	ABN96843 standard, DNA; 2173 BP.	
XX		
AC	ABN96843;	
XX		
DT	13-AUG-2002 (first entry)	
XX		
DE	Gene #3341 used to diagnose liver cancer.	
XX		
KW	Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;	
KW	metastatic liver tumor; cytosolic; expression profile; disease state;	
KW	disease progression; drug toxicity; drug efficacy; drug metabolism.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200229103-A2.	
XX		
PD	11-APR-2002.	
XX		
PF	02-OCT-2001; 2001WO-US30589.	
XX		
PR	02-OCT-2000; 2000US-237054P.	
XX		

PA	(GENE-) GENE LOGIC INC.
PI	Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
XX	WPI; 2002-426119/45.
DR	
XX	
PT	Diagnosing and detecting the progression of liver cancer,
PT	hepatocellular carcinoma or metastatic liver tumor in a patient,
PT	involves detecting the level of expression of two or more genes in a
PT	liver tissue sample -
XX	
P5	Claim 1; SEQ ID NO 3341; 298bp; English.
XX	
CC	The invention relates to a novel method for diagnosing and detecting the
CC	progression of liver cancer, hepatocellular carcinoma or metastatic live
CC	tumour in a patient, and differentiating metastatic liver cancer from
CC	hepatocellular carcinoma in a patient, involving detecting the level of
CC	expression of two or more genes represented in ABN97455 in a
CC	tissue sample. The method of the invention has hepatotropic, and
CC	cytostatic activity. The method is useful for diagnosing and detecting
CC	the progression of liver cancer, hepatocellular carcinoma and metastatic
CC	liver carcinoma in a patient. The method is useful for identifying
CC	expression profiles which serve as useful diagnostic markers as well as
CC	markers that can be used to monitor disease stages, disease progression,
CC	drug toxicity, drug efficacy and drug metabolism.
CC	Note: The sequence data for this patent did not form part of the printed
CC	specification, but was obtained in electronic format directly from WIPO
CC	at ftp.wipo.int/pub/published_pct_sequences.
XX	
SQ	Sequence 2173 BP; 606 A; 431 C; 503 G; 633 T; 0 other;
	Alignment Scores:
	Pred. No.: 4.29e-86 Length: 2173
	Score: 847.50 Matches: 162
	Percent Similarity: 97.63% Conservative: 3
	Best Local Similarity: 95.86% Mismatches: 3
	Query Match: 81.26% Indels: 1
	DB: 24 Gaps: 1
	US-09-782-953-7 (1-198) x ABN96843 (1-2173)
OY	30 AlaIysPhegiUsErLeuPheArgThrTyAspLyAspThrThrphegIntyrPhelys 49
Db	58 GCCAAATTGGAGTCCCTTTAGACGTATGACAAGACATCACCTTCAGTATTTTAAG 117
OY	50 SerPheLySArGVAlArgLIeaSnPheSerAnProLeuSerAlaLaApAlaArGVleu 69
Db	118 AGCTCAACAGCGTAGATAAATCTCAGCAACCCTTCGCCAGAGATGCAGGCTC 177
OY	70 ArgLeuHisLylsThngInuPheLeuGIlyLyAgILuMetLyIsLeuTyRheAlagIntHleu 89
Db	178 CAGCGCATTAAGACGATTTCTGGGAAGAAGAAATGAAAGTTATATTTTGCTCAGACTTA 237
OY	90 HisLIeGySeSrHiLeuAlAProProAbnProAspLySGInPheLeuIIseerPro 109
Db	238 CACATAGGAAGCTCACACTGCTGCCCAAATCCAACAAGCAGTTTCTGATCTCCCCCT 257
OY	110 ProAlaserProProValGIlyTrpLySGInuVAlguApAlathrProValIIleaenTyR 129
Db	298 CCCGCTCTCCGCCAGTGGATGGAAACAAGTGAAGATGGAGACCCAGTCATPAACTAT 357
OY	130 AppLeuLeuTYrAlaIIseerLyLeuGIlyProGLytluluyrtyrgIuleuHIlaAla 149
Db	358 GATCTCTTAATNGCATCTCCAACTGGGGCCAGGGGAAAAGTATGCAATTGCACGCGCG 417
OY	150 ThrAspProThrProSerValIVAlValHlSVAGIySGISeRaSpGINuIAenGIuGLu 169
Db	418 ACTGACACCACTCCAGGGGTGGTGTCCATTAATGATGAGTATCAAGAG--AAGGAG 474
OY	170 GIUGIUGIUGIuMeGIuArMelLySArGProluysProluysIleIlegIntHrArgx 189
Db	475 GAAGAAGAGGAATGGAAGAATGAGAGACCTTAAGCCAAATAATTATCCAGACGAGAGG 534

QY 190 ProGluTyrThrProIleHisLeuSer 198
 |||||
 Db 535 CCGGAGTACAGCCGATCCACCTCAGC 561

RESULT 8

AAAX01282
 ID AAX01282 standard; cDNA to mRNA; 2174 BP.

XX
 AC AAX01282;

DT 09-APR-1999 (first entry)

DE Human DSCR1 coding sequence.

XX DSCR1; Down's Syndrome critical region 1; human; pathogenic abnormality;
 KW Central Nervous System development; mental retardation; heart defect; ds.

XX Homo sapiens.

FH Key Location/Qualifiers

FT CDS 49..564

FT /tag= a

FT polyA_signal 1541..1546

FT /tag= b

FT polyA_signal 2132..2137

FT /tag= c

XX US5869318-A.

XX 09-FEB-1999.

XX 07-JUN-1996; 96US-0665040.

XX 07-JUN-1995; 95ES-0001140.

XX (PAL/) PALLEJA X E.

XX Fuentes JJ, Palleja XE, Pritchard M;

XX WPI; 1999-152781/13.

XX P-PSDB; AAW73898.

XX DNA encoding foetal brain proteins - believed to be associated with
 PT Down's syndrome

PS Claim 1; Column 15-18; 19pp; English.

XX This sequence is the Down's Syndrome critical region 1 (DSCR1) gene
 CC of the invention. The DSCR1 gene was found to be located in the
 CC q22.1-22.2 region of human chromosome 21. An increase in the transient
 CC expression of DSCR1 mRNA in the brains of young rats, compared to
 CC expression levels in the brains of adult rats, suggests an important role
 CC for DSCR1 during the development of the Central Nervous System (CNS), and
 CC that over expression of DSCR1 may be involved in pathogenic abnormalities
 CC of mental retardation and/or heart defects as found in Down's syndrome
 CC patients.

XX SQ Sequence 2174 BP; 605 A; 436 C; 502 G; 628 T; 3 other;

Alignment Scores:

Pred. No.: 9,4e-86 Length: 2174
 Score: 844.50 Matches: 161
 Percent Similarity: 97.63% Conservative: 4
 Best Local Similarity: 95.27% Mismatches: 3
 Query Match: 80.97% Indels: 1
 DB: 20 Gaps: 1

US-09-782-953-7 (1-198) x AAX01282 (1-2174)

QY 30 AlaLysPheGluSerLeuPheArgThrTyAspLysAspThrPheGlnTyrPheLys 49
 |||||
 Db 58 GCCAAATTTGAGTCCCTTTAGGACGTATGACAGGACATCACCTTTTTCAGTATTITTAAG 117

QY 50 SerPheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAspAlaArgLeu 69
 |||||
 Db 118 AGCTTCAAACGAGTCAAGATAAACTTCAGCAACCCCTTCTCCGACAGCATGCCAGGCTC 177

QY 70 ArgLeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyrPheAlaGlnThrLeu 89
 |||||
 Db 178 CAGCTGCATAAGACTGAGTTTCTGGGAAAGGAAATGAAGTTATATTTTCTCAGACCTTTA 237

QY 90 HisIleGlySerSerHisLeuAlaProProAsnProAspLysGlnPheLeuIleSerPro 109
 |||||
 Db 238 CACATAGGAAGCTCACACCTGGCTCCGCCAATCCAGACACAGCAGTTTCTGATCTCCCT 297

QY 110 ProAlaSerProProValGlyTrpLysGlnValGluAspAlaThrProValIleAsnTyr 129
 |||||
 Db 298 CCCGCTCTCCGCCAGTGGATGGAACAACATGGAAGATGCCACCCAGTCATAAATAT 357

QY 130 AspLeuLeuTyrAlaIleSerLysLeuGlyProGlyGluLysTyrGluLeuHisAlaAla 149
 |||||
 Db 358 GATCTCTTATATGCCATCTCCAAAGCTGGGGCCAGGGAAAGATATGAATTCACCGCAGC 417

QY 150 ThrAspProThrProSerValValHisValHisValCysGluSerAspGlnGluAsnGlu 169
 |||||
 Db 418 ACTGACACCACTCCACAGCGTGGTGGTCCATGTATGTGAGAGTGCATCAAGAG 474

QY 170 GluGluGluGluMetGluArgMetLysArgProLysProLysIleIleGlnThrArgArg 189
 |||||
 Db 475 GAAGAAGAGGAAATGGAAGATGAGGAGACCTAAGCCAAAATTTATCCAGACCAGGAGG 534

QY 190 ProGluTyrThrProIleHisLeuSer 198
 |||||
 Db 535 CCGGAGTACAGCCGATCCACCTCAGC 561

RESULT 9

AAAF18328

ID AAF18328 standard; DNA; 2358 BP.

XX AAF18328;

DT 14-MAR-2001 (first entry)

DE Lung cancer associated polynucleotide sequence SEQ ID 347.

XX Human; lung cancer associated protein; neuroprotective; cytostatic;
 KW cardioactive; immunomodulatory; muscular active; vulnerary;
 KW gastrointestinal; nephrotropic; antiinfective; gynecological;
 KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
 KW proliferative disorder; wound healing; infectious disease; ds.

XX Homo sapiens.

XX WO200055180-A2.

XX 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US05918.

XX 12-MAR-1999; 99US-0124270.

XX (HUMA-) HUMAN GENOME SCI INC.

XX (ROSE/) ROSEN C A.

XX Ruben SM;

XX WPI; 2000-587514/55.

XX P-PSDB; AAB58452.

XX Lung cancer associated gene sequences, referred to as lung cancer
 PT antigens, useful for treatment, prevention, and diagnosis of disorders
 PT such as lung cancer -

XX Claim 1; Page 805-806; 1425pp; English.

XX Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer

Db 614 GTGACGCTGTGACAGTACATAGAGAGAGAGACCAAG----- 658
 Oy 178 LysArgProLysProLysIleIleGlnThrArgArgProGluTyrThrPro 194
 Db 659 ACTTCCCAAGCCAAATATCATCTCAAACTCGGCGCTCGGCTGCCACCC 709

RESULT 12
 ABA91457
 ID ABA91457 standard; cDNA; 3159 BP.
 AC ABA91457;
 XX
 DT 18-APR-2002 (first entry)
 XX
 DE Human Down syndrome critical region 1-like 1 protein cDNA.
 KM Down syndrome critical region 1-like 1; DSCR1L1 alpha; human;
 KM Down syndrome; Alzheimer's disease; dementia; transgenic;
 KM neuroprotective; noctropic; anticonvulsant; diagnosis;
 KM gene therapy; gene; 88.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 185..952
 FT /*tag= a
 FT /product= "DSCR1L1"
 FT /transl_except= (pos:215..217, aa:Xaa)
 FT /note= "Xaa = unknown"
 FT
 PN WO200204513-A2.
 PD 17-JAN-2002.
 XX
 PF 11-JUL-2001; 2001WO-US21982.
 XX
 PR 11-JUL-2000; 2000US-0614474.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Loring JF, Tingley DM, Edwards CM, Streeter DG;
 DR WPI; 2002-164633/21.
 DR P-PSDB; AAM50760.
 XX
 PT Novel Down syndrome critical region 1-like 1 protein and nucleic acid
 PT encoding the protein useful for diagnosis and treatment of Alzheimer's
 PT disease, Down syndrome and other forms of dementia -
 PS
 PS Claim 2; Page 46-47; 54pp; English.

The present sequence is that of cDNA clone incyte ID No: 247500.5
 encoding novel down syndrome critical region 1-like 1 protein
 (DSCR1L1 alpha, see AAM50760). Northern analysis indicated
 expression of DSCR1L1 alpha in various libraries, with the highest
 abundance in tissues from the nervous system, including tissues
 associated with schizophrenia, Huntington's disease, epilepsy and
 amyotrophic lateral sclerosis. An absence of DSCR1L1 expression
 was observed in 7 of 8 libraries from subjects with Alzheimer's
 disease. A claimed method for detecting differential expression of
 a nucleic acid encoding DSCR1L1 alpha can be used to diagnose Down
 CC syndrome, Alzheimer's disease and other forms of dementia. A probe
 CC from such a nucleic acid is useful for identifying naturally
 CC occurring molecules encoding DSCR1L1 alpha, allelic variants or
 CC other molecules. The nucleic acids are useful for producing
 CC transgenic cell lines or organisms which model human disorders.
 CC They may also be used in gene therapy, and to screen for ligands
 CC which specifically bind the nucleic acid molecule, such as
 CC a DNA or RNA molecule, peptide nucleic acid, artificial chromosome,
 CC peptide, transcription factor, repressor or regulatory molecule,
 CC for use as a therapeutic.
 CC
 XX Sequence 3159 BP; 926 A; 657 C; 657 G; 915 T; 4 other;

Alignment Scores:
 Pred. No.: 2,266-60 Length: 3159
 Score: 623.00 Matches: 123
 Percent Similarity: 73.10% Conservative: 21
 Best Local Similarity: 62.44% Mismatches: 41
 Query Match: 59.73% Indels: 12
 DB: 24 Gaps: 3

US-09-782-953-7 (1-198) x ABA91457 (1-3159)

Oy 2 AspPheArgAspPheSerTyrAsnPheSerSerLeuIleAlaCysValAlaAsnAsp 21
 Db 371 GACTTCATGACCTCTCC-----AACTCGTGTGTTGGTGCAATGTTCACTGCA 421
 Oy 22 ValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgTyrAspLys 41
 Db 422 GTGTTTGAG 481
 Oy 42 AspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsnPro 61
 Db 482 TGTGTGACGTTCCAGCTATTATTAAGAGTTTCAGACGTGTCCTATMAACTTCAGCAATCCT 541
 Oy 62 LeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGluMet 81
 Db 542 AAATGTGACGCCCGAGCTAGATGAGCTTATGAAACCAATTCAGAGGAGAGAGAGAGAG 601
 Oy 82 LysLeuTyrPheAlaGlnThrLeuHisIleGlySerSer-----HisLeuAla 97
 Db 602 AAGCTCTACTTTGGACAGGTTCCAGCTCCAGACAGACAGATGAGAGAGAGAGAGAGAGAG 661
 Oy 98 ProProAsnProAspLysGlnPheLeuIleSerProProAlaSerProProValGlyTyr 117
 Db 662 CCACCCCGACCTGCGCAACAGTTTCTATCTCGCCCTCTCTCCCACTGTTGGCTGG 721
 Oy 118 LysGlnValGluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLys 137
 Db 722 CAGCCCATCAACGATGCCACGCCGAGCTCTCAACTATGACTCTCTATGCTGTGGCCAA 781
 Oy 138 LeuGlyProGlyGluLysTyrGlnLeuHisAlaIleThrAspProThrProSerValVal 157
 Db 782 CTAGGACCGAGGAGAGAGAGATGATGAGCTCCATGACAGGAGCTGACCCCAAGTGTCTG 841
 Oy 158 ValHisValCysGluSerAspGlnGluAsnGluGluGluGluGluGluMet 177
 Db 842 GTGACGCTGTGACAGACATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 886
 Oy 178 LysArgProLysProLysIleIleGlnThrArgArgProGluTyrThrPro 194
 Db 887 ACTTCCCAAGCCAAATATCATCTCAAACTCGGCGCTCGGCTGCCACCC 937

RESULT 13
 AAD30157
 ID AAD30157 standard; DNA; 3184 BP.
 XX
 AC AAD30157;
 XX
 DT 17-MAY-2002 (first entry)
 XX
 DE Human MCIP associated DNA #3.
 XX
 KM Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;
 KM heart failure; cardiomyopathy; heart disease; human; gene; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 205..783
 FT /*tag= a
 FT /product= "Human MCIP associated protein #3"
 FT
 PN WO200204491-A2.

PD 17-JAN-2002.
 XX
 PF
 XX 06-JUL-2001; 2001WO-US21662.
 XX
 PR 07-JUL-2000; 2000US-216601P.
 PR 13-FEB-2001; 2001US-0782953.
 XX
 PA (TEXA) UNIV TEXAS SYSTEM.
 PA (WILL/) WILLIAMS S R.
 PA (ROTH/) ROTHERMEL B.
 XX
 XX Williams SR, Rothermel B;
 XX MPI: 2002-179698/23.
 DR P-PSDB; AAE18916.
 DR
 XX
 PT Screening for modulators of muscle calcineurin interacting protein
 PT (MCIP) binding, expression or phosphorylation, useful for treating
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,
 PT calcineurin and a test compound -
 XX
 PS Disclosure; Page 163-165; 174pp; English.
 XX
 CC The invention relates to muscle calcineurin interacting proteins (MCIPs)
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical
 CC complex with the catalytic subunit of calcineurin and increased levels
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate
 CC transcription of certain target genes. The invention also relates to
 CC methods for identifying modulators of MCIP binding, expression or
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin
 CC may be used for treating cardiac hypertrophy and heart failure.
 CC Antibodies to MCIP can be used in characterizing the MCIP content of
 CC healthy and diseased tissues and subsequently for determining the
 CC presence or absence of cardiomyopathy or as predictor of heart disease.
 CC The present sequence is human MCIP associated DNA.
 CC Note: This sequence has been described as human MCIP3 in the
 CC specification, however the sequence seems to be a polynucleotide encoding
 CC a MCIP associated protein.
 XX
 SQ Sequence 3184 BP; 921 A; 681 C; 657 G; 925 T; 0 other;
 Alignment Scores:
 Pred. No.: 1.63e-59 Length: 3184
 Score: 615.50 Matches: 119
 Percent Similarity: 75.53% Conservative: 23
 Best Local Similarity: 63.30% Mismatches: 37
 Query Match: 59.01% Indels: 9
 DB: 24 Gaps: 2
 US-09-782-953-7 (1-198) x AAD30157 (1-3184)
 QY 11 SerSerLeuIleAlaCysValAlaAsnAspValPheSerGluSerGluThrArgAla 30
 DB 220 TCCACTCTGTTGCTGCTGGTGGATGTCAGGCTTTTACCAATCAGGAGGTTAAGGAA 279
 QY 31 LysPheGluSerLeuPheArgThrThrAspLysAspThrPheGlnTyrPheLysSer 50
 DB 280 AATTTGGGGGAGCTGTTTCGAGCTTATGATGACTGTGTGACGTTCCAGCTATTTAAGCT 339
 QY 51 PheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAspAlaArgLeuArg 70
 DB 340 TTCAGAGCTGTCGTATTAACCTTACCAATCTTAATCTGACCCGAGCTAGGATAGAG 399
 QY 71 LeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyrPheAlaGlnThrLeuHis 90
 DB 400 CTTTCATGAAACCAATTCAGAGGGAATAAATAAGCTCTACTTTGCACAGTTTCAGCT 459
 QY 91 IleGlySerSer-----HisLeuAlaProProAsnProAspLysGlnPheLeu 106
 DB 460 CCAGACAGATGGAGACAAATGCGCTTGGCTCCACCCAGCTGCCAAACAGTTTCTC 519
 QY 107 IleSerProProAlaSerProProValGlyTyrPheGlnValGluAspAlaThrProVal 126
 DB 02-OCT-2000; 2000US-237173P. 02-OCT-2000; 2000US-237278P.

DB 520 ATCTGCCCTTCCTCCACCTGTTAGTGGCAGCCCATCAACGATGCCAGCCAGTC 579
 QY 127 IleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyProGlyGluLysTyrGluLeu 146
 DB 580 CTCAACTATGACCTCTCTATGCTGTGGCCAACTAGGACCAAGAGAGATATGAGCTC 639
 QY 147 HisAlaAlaThrAspProThrProSerValValHisValCysGluSerAspGlnGlu 166
 DB 640 CATGCAGGAGCTGAGTCCACCCCAAGTGTGCTGCGACGTCGCGACAGTGACATAGAG 699
 QY 167 AsnGluGluGluGluGluMetGluArgMetLysArgProLysProLysIleLeuGln 186
 DB 700 GAAGAAGAGAGCCCAAG-----ACTTCCCAAGCAAAATCATCCAA 744
 QY 187 ThrArgArgProGluTyrThrPro 194
 DB 745 ATCGGGCTCTGGCTGCCACCC 768
 RESULT 14
 ABL61768
 ID ABL61768 standard; DNA; 3184 BP.
 XX
 AC ABL61768;
 XX
 DT 15-MAY-2002 (first entry)
 XX
 DE Colon adenocarcinoma related gene sequence SEQ ID NO:105.
 KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
 KW cytostatic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;
 KW gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200194629-A2.
 XX
 PD 13-DEC-2001.
 XX
 PF 30-MAY-2001; 2001WO-US10838.
 XX
 PR 05-JUN-2000; 2000US-209473P.
 PR 05-JUN-2000; 2000US-209531P.
 PR 18-SEP-2000; 2000US-233133P.
 PR 18-SEP-2000; 2000US-233617P.
 PR 20-SEP-2000; 2000US-234009P.
 PR 20-SEP-2000; 2000US-234034P.
 PR 20-SEP-2000; 2000US-234052P.
 PR 22-SEP-2000; 2000US-234509P.
 PR 22-SEP-2000; 2000US-234567P.
 PR 25-SEP-2000; 2000US-234923P.
 PR 25-SEP-2000; 2000US-234924P.
 PR 25-SEP-2000; 2000US-235077P.
 PR 25-SEP-2000; 2000US-235082P.
 PR 25-SEP-2000; 2000US-235134P.
 PR 26-SEP-2000; 2000US-235637P.
 PR 26-SEP-2000; 2000US-235638P.
 PR 27-SEP-2000; 2000US-235711P.
 PR 27-SEP-2000; 2000US-235720P.
 PR 27-SEP-2000; 2000US-235840P.
 PR 27-SEP-2000; 2000US-235863P.
 PR 28-SEP-2000; 2000US-236028P.
 PR 28-SEP-2000; 2000US-236032P.
 PR 28-SEP-2000; 2000US-236033P.
 PR 28-SEP-2000; 2000US-236034P.
 PR 28-SEP-2000; 2000US-236109P.
 PR 28-SEP-2000; 2000US-236111P.
 PR 29-SEP-2000; 2000US-236842P.
 PR 29-SEP-2000; 2000US-236891P.
 PR 02-OCT-2000; 2000US-237172P.
 PR 02-OCT-2000; 2000US-237173P.
 PR 02-OCT-2000; 2000US-237278P.

```

PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.
PA (AVALON PHARM.
XX
XX
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppe DR, Weaver Z;
XX
XX WPI; 2002-188264/24.
XX
XX Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set -
XX
XX Claim 1; SEQ ID 105; 44pp; English.
XX
XX The present invention describes a method (M1) for screening for an
XX anti-neoplastic agent. The method involves exposing cells to a chemical
XX agent to be tested for anti-neoplastic activity, determining a change in
XX expression of at least one gene (1) of a signature gene set, where (1)
XX comprises a sequence (S) selected from 8447 sequences (given in ABL61664
XX to ABL70110), or is at least 95% identical to (S), where a change in
XX expression is indicative of anti-neoplastic activity. (1) has cytostatic
XX activity and can be used in gene therapy. M1 can be used for screening
XX an anti-neoplastic agent, and can be used for producing a product which
XX is the data collected with respect to the anti-neoplastic agent as a
XX result of M1, and the data is sufficient to convey the chemical
XX structure and/or properties of the agent. M1 can be used in the
XX treatment of cancer such as colon, breast, stomach, lung, thyroid,
XX oesophageal, ovarian, kidney, prostate or pancreatic cancer,
XX adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
XX infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
XX carcinoma, papillary carcinoma and Wilms' tumour.
XX
XX Sequence 3184 BP; 921 A; 681 C; 657 G; 925 T; 0 other;
SQ

```

Alignment Scores: 1.63e-59 Length: 3184
Pred. No.: 615.50 Matches: 119
Score: 75.53% Conservative: 23
Best Local Similarity: 63.30% Mismatches: 37
Query Match: 59.01% Indels: 9
DB: 24 Gaps: 2

```

US-09-782-953-7 (1-198) x ABL61768 (1-3184)
OY 11 SerSerLeuilealacysvalalaaaspaaspvalpheserluserglutrhargala 30
DB 220 TCACACTGGTTCCTGCTGTGCGATGTCGAGGTCTTTACCAATGACGAGGTAAAGAA 279
OY 31 LysPheGluSerLeuPheAlaGlyThrTyrAspLysAspThrThrPheGlnTyrPheLysSer 50
DB 280 AAATTGGGGGAGCTGTTTCGACTTATGATGACTGTGTGACGTTCCAGCTATTAAAGAGT 339
OY 51 PheLysArgValArgIleAsnPheSerAspProLeuSerAlaIleAspAlaArgLeuArg 70
DB 340 TTGAGAGGTGTCGTAATTAACCTTCAGCAATCTTAATCTCAGCCGACGTAGATAGAG 399
OY 71 LeuHieLysThrGluPheLeuGlyLysGluMetLysLeuTyrPheAlaGlnThrLeuHis 90
DB 400 CTTCATGAAACCCCAATTCAGAGGGAATAAATTAAGCTTACCTTGACAGGTTCAAGACT 459
OY 91 IleGlySerSer-----HisLeuAlaProProAsnProAspLysGlnPheLeu 106
DB 460 CCAGAGACAGATGGAGCAAACTGCACTTGCTCCACCCGCTGCCAAACAGTTTCTC 519

```

```

OY 107 IleSerProBlaAserProProValGlyTyrLysGlnValGluAspAlaThrProVal 126
DB 520 ATTCGCCCCCTTCCTCCACCTGTTAGCTGGAGCCCATCAACGATGCGACGCCAGTC 579
OY 127 IleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlnTyrProGlyGlyLysTyrGluLeu 146
DB 580 CTCAACTATGACCTCCCTCTGCTGTGGCCAAACTAGACAGAGAGAAATGATGAGCTC 639
OY 147 HisAlaIleThrAspProThrProSerValValHisValCysGluSerAspGlnGlu 166
DB 640 CATCAGGAGACTGAGTGCATCCACCACCAAGTGTCTGTGCACTGTCCAGCATGATGAG 699
OY 167 AsnGluGluGluGluGluGluMetLysArgMetLysArgProLysProLysIleIleGln 186
DB 700 GAAGAGAGAGACCAAG-----ACTTCCCAAGCCAAATAATCATCAA 744
OY 187 ThrArgArgProGluTyrThrPro 194
DB 745 ACTGGGCTCTGCGCTGCCACCC 768

```

RESULT 15
ABL65187
ID ABL65187 standard; DNA, 3184 BP.
XX
XX ABL65187;
AC
XX
XX 15-MAY-2002 (first entry)
DT
XX
XX Lung cancer related gene sequence SEQ ID NO:3524.
DE
XX
XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
XX stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
XX cytostatic; gene therapy; anti-neoplastic; Wilms' tumour; adenocarcinoma;
XX gene; ds.
XX
XX Homo sapiens.
OS
XX
XX WO200194629-A2.
PN
XX
XX 13-DEC-2001.
PD
XX
XX
XX 30-MAY-2001; 2001WO-US10838.
PF
XX
XX 05-JUN-2000; 2000US-209473P.
XX
XX 05-JUN-2000; 2000US-209531P.
XX
XX 18-SEP-2000; 2000US-231333P.
XX
XX 18-SEP-2000; 2000US-233617P.
XX
XX 20-SEP-2000; 2000US-234009P.
XX
XX 20-SEP-2000; 2000US-234034P.
XX
XX 20-SEP-2000; 2000US-234052P.
XX
XX 22-SEP-2000; 2000US-234509P.
XX
XX 22-SEP-2000; 2000US-234567P.
XX
XX 25-SEP-2000; 2000US-234923P.
XX
XX 25-SEP-2000; 2000US-234924P.
XX
XX 25-SEP-2000; 2000US-235077P.
XX
XX 25-SEP-2000; 2000US-235082P.
XX
XX 25-SEP-2000; 2000US-235134P.
XX
XX 25-SEP-2000; 2000US-235280P.
XX
XX 26-SEP-2000; 2000US-235637P.
XX
XX 26-SEP-2000; 2000US-235638P.
XX
XX 26-SEP-2000; 2000US-235711P.
XX
XX 27-SEP-2000; 2000US-235720P.
XX
XX 27-SEP-2000; 2000US-235840P.
XX
XX 27-SEP-2000; 2000US-235863P.
XX
XX 28-SEP-2000; 2000US-236028P.
XX
XX 28-SEP-2000; 2000US-236032P.
XX
XX 28-SEP-2000; 2000US-236033P.
XX
XX 28-SEP-2000; 2000US-236034P.
XX
XX 28-SEP-2000; 2000US-236109P.
XX
XX 28-SEP-2000; 2000US-236111P.
XX
XX 29-SEP-2000; 2000US-236842P.
XX
XX 29-SEP-2000; 2000US-236891P.

PR 02-OCT-2000; 2000US-237172P.
 PR 02-OCT-2000; 2000US-237173P.
 PR 02-OCT-2000; 2000US-237278P.
 PR 02-OCT-2000; 2000US-237294P.
 PR 02-OCT-2000; 2000US-237295P.
 PR 02-OCT-2000; 2000US-237316P.
 PR 03-OCT-2000; 2000US-237425P.
 PR 03-OCT-2000; 2000US-237598P.
 PR 03-OCT-2000; 2000US-237604P.
 PR 03-OCT-2000; 2000US-237606P.
 PR 03-OCT-2000; 2000US-237608P.
 PR 01-NOV-2000; 2000US-244867P.
 PR 01-NOV-2000; 2000US-245084P.
 XX
 PA (AVAL-) AVALON PHARM.
 XX
 PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 PI Soppet DR, Weaver Z;
 XX
 DR WPI; 2002-188264/24.
 XX

PT Screening for anti-neoplastic agent involves exposing cells to a
 PT chemical agent to be tested for anti-neoplastic activity, and
 PT determining a change in expression of a gene of a signature gene set
 XX
 PS Claim 1; SEQ ID 3524; 44pp; English.
 XX

CC The present invention describes a method (M1) for screening for an
 CC anti-neoplastic agent. The method involves exposing cells to a chemical
 CC agent to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 847 sequences (given in ABL61664
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening
 CC an anti-neoplastic agent, and can be used for producing a product which
 CC is the data collected with respect to the anti-neoplastic agent as a
 CC result of M1, and the data is sufficient to convey the chemical
 CC structure and/or properties of the agent. M1 can be used in the
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilms' tumour.
 XX

SQ Sequence 3184 BP; 921 A; 681 C; 657 G; 925 T; 0 other;

Alignment Scores:

Pred. No.:	1-63e-59	Length:	3184
Score:	615.50	Matches:	119
Percent Similarity:	75.53%	Conservative:	23
Best Local Similarity:	63.30%	Mismatches:	37
Query Match:	59.01%	Indels:	9
DB:	24	Gaps:	2

US-09-782-953-7 (1-198) x ABL65187 (1-3184)

QY 11 SerSerLeuLeaCysValAlaAenAspValPheSerGluThrArgAla 30
 Db 220 TCCACTCTGGTTCCTGCTGGTGGATGTCGAGTCTTTACCAATCAGGAGGTTAAGGAA 279
 QY 31 LysPheGluSerLeuPheArgThrTyAspLysAspThrPheGlnTyrPheLysSer 50
 Db 280 AATTTGGGGGACTGTTTCGGACTTTATGACTGTGTGACGTTCCAGCTATTTTAAGAGT 339
 QY 51 PheLysArgValArgLeuAenPheSerAsnProLeuSerAlaAlaAspAlaArgLeuArg 70
 Db 340 TTCAGACGTCGCTAACTTACCAATCTTAATCTGCAGCCCGAGCTAGGATAGAG 399
 QY 71 LeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyrPheAlaGlnThrLeuHis 90
 Db 400 CTTTCATGAACCAATTCAGAGGGAATAAAGCTCTACTTTGCACAGGTTTCAGACT 459

QY 91 IleGlySerSer-----HisLeuAlaProProAsnProAspLysGlnPheLeu 106
 Db 460 CCAGAGACAGATGGAGACAAACTGGCTCCACCCAGCCTGCCAAACAGTTTCTC 519
 QY 107 IleSerProProAlaSerProProValGlyTTPlyGlnValGluAspAlaThrProVal 126
 Db 520 ATCTCGCCCTTCTCTCCACCTGTTAGCTGCAGCCCATCAACGATGCCAGCCAGTC 579
 QY 127 IleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyProGlyGluLysTyrGluLeu 146
 Db 580 CTCAACTATGACCTCTCTATGCTGGCCAAACTAGGACCCAGAGAGATGATGAGCTC 639
 QY 147 HisAlaAlaThrAspProThrProSerValValHisValCysGluSerAspGlnGlu 166
 Db 640 CATCGAGGACTGAGTCCACCCCAAGTGTGTCACGTCGTCGACAGTGCACATAGAG 699
 QY 167 AsnGluGluGluGluGluMetGluArgMetLysArgProLysProLysIlelleGln 186
 Db 700 GAAGAAGAGGCCCAAG-----ACTTCCCAAGCCAAAAATCATCCAA 744
 QY 187 ThrArgArgProGluTyrThrPro 194
 Db 745 ACTCGGCGTCTGGCCTGCCACCC 768

Search completed: December 14, 2002, 20:09:05
 Job time : 145.568 secs

THIS PAGE BLANK (uspto)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2002, 14:11:25 ; Search time 1079.63 Seconds
(without alignments)
16146.861 Million cell updates/sec

Title: US-09-782-953-1
Perfect score: 599
Sequence: 1 gagggcgaagaagacctcca.....gaggagggaaggaagagat 599

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.on.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pln.*
- 35: em.htg.rod.*
- 36: em.htg.mam.*
- 37: em.htg.vrt.*
- 38: em.sv.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	599	100.0	599	6	AX365311	AX365311 Sequence
2	573	95.7	2198	10	BC013551	BC013551 Mus muscu
3	524.4	87.5	2141	10	AF263239	AF263239 Mus muscu
4	504.4	84.2	597	6	AX365315	AX365315 Sequence
5	504.4	84.2	597	10	AF237790	AF237790 Mus muscu
6	453.2	75.7	2216	10	CGU60263	U60263 Cricetus
7	435.8	72.8	626	10	AB075973	AB075973 Rattus no
8	427.8	71.4	621	10	AF282255	AF282255 Mus muscu
9	427.8	71.4	2125	10	AF260717	AF260717 Mus muscu
10	426.2	71.2	597	6	AX365312	AX365312 Sequence
11	426.2	71.2	597	10	AF237789	AF237789 Mus muscu
12	425.4	71.0	2224	10	AF263240	AF263240 Mus muscu
13	413.8	69.1	2331	6	AX365321	AX365321 Sequence
14	413.8	69.1	2346	9	HSU85267	HSU85267 Homo sapien
15	404.8	67.6	2348	6	AX281651	AX281651 Sequence
16	394	65.8	562	9	HSU53821	U53821 Homo sapien
17	349.2	58.3	2212	6	AX365324	AX365324 Sequence
18	349.2	58.3	2227	9	HSU85266	U85266 Homo sapien
19	349.2	58.3	2289	9	BC002864	BC002864 Homo sapi
20	349.2	58.3	2407	9	AK092184	AK092184 Homo sapi
21	343.8	57.4	2284	9	HSU85265	U85265 Homo sapien
22	341.8	57.1	2173	6	AX410694	AX410694 Sequence
23	341.8	57.1	2173	9	HSU28833	U28833 Homo sapien
24	341.2	57.0	2679	9	AK055845	AK055845 Homo sapi
25	340.2	56.8	798	9	AF400429	AF400429 Homo sapi
26	340.2	56.8	2174	6	AR034241	AR034241 Sequence
27	202.2	33.8	3184	6	AX329596	AX329596 Sequence
28	202.2	33.8	3184	6	AX333015	AX333015 Sequence
29	202.2	33.8	3184	6	AX365327	AX365327 Sequence
30	202.2	33.8	3184	6	AX420436	AX420436 Sequence
31	202.2	33.8	3184	9	HUM2AK14	D33407 ZAK1-4 mRNA
32	197.6	33.0	934	6	AX074351	AX074351 Sequence
33	197.6	33.0	3159	6	AX420425	AX420425 Sequence
34	197.6	33.0	3261	9	AK090990	AK090990 Homo sapi
35	194.4	32.5	3240	9	AY034085	AY034085 Homo sapi
36	194.4	32.5	3253	9	AY034086	AY034086 Homo sapi
37	191	31.9	594	6	AX365318	AX365318 Sequence
38	191	31.9	594	10	AF237791	AF237791 Mus muscu
39	191	31.9	659	10	AF237887	AF237887 Mus muscu
40	189.8	31.7	720	6	AX365333	AX365333 Sequence
41	189.8	31.7	720	9	AF176117	AF176117 Homo sapi
42	189.4	31.6	3227	10	AB061524	AB061524 Mus muscu
43	188.4	31.5	828	6	AX365330	AX365330 Sequence
44	188.4	31.5	828	6	AX420437	AX420437 Sequence
45	188.4	31.5	828	9	AF176116	AF176116 Homo sapi

ALIGNMENTS

RESULT 1
AX365311
LOCUS AX365311
DEFINITION Sequence 1 from Patent WO0204491.
ACCESSION AX365311
VERSION AX365311.1 GI:18697042
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Williams, S.R. and Rothermel, B.
TITLE Methods and compositions relating to muscle selective calcineurin interacting protein (mcip)

JOURNAL Patent: WO 0204491-A 1 17-JAN-2002;
 Board of Regents, The University of Texas System (US); Williams,
 Sanders R. (US); Rochester, Beverly (US)

FEATURES
 source Location/Qualifiers
 1.599
 /organism="Mus musculus"
 /db_xref="taxon:10090"

BASE COUNT 164 a 150 c 149 g 136 t

ORIGIN

Query Match 100.0%; Score 599; DB 6; Length 599;
 Best Local Similarity 100.0%; Pred. No. 1,46-179;
 Matches 599; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGGGCAAGAGAACTCCAGCTTGGCTTGAAGAGAGAGAGCTCTTCCGTTAAGCT 60
 DB 1 GAGGTGCAAGAGAACTCCAGCTTGGCTTGAAGAGAGAGAGCTCTTCCGTTAAGCT 60
 QY 61 CTGCCCCGTAAGAAAGCAGATGATTTTAGGAGCTTTAGCTCAATTTTGGCTCCCTGAT 120
 DB 61 CTGCCCCGTAAGAAAGCAGATGATTTTAGGAGCTTTAGCTCAATTTTGGCTCCCTGAT 120
 QY 121 TGGTGTGTGGCAAGAGATGATCTTTCAGCCAAAGTGAAGAGAGAGAGAGAGAGAG 180
 DB 121 TGGTGTGTGGCAAGAGATGATCTTTCAGCCAAAGTGAAGAGAGAGAGAGAGAGAG 180
 QY 181 CCTCTTCAGAAATATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
 DB 181 CCTCTTCAGAAATATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
 QY 241 CCGGATTAACCTTGAGCAACCCCTTATCTGAGCCGATGCCAGCTGCGCTGCACAGAC 300
 DB 241 CCGGATTAACCTTGAGCAACCCCTTATCTGAGCCGATGCCAGCTGCGCTGCACAGAC 300
 QY 301 CGAGTTCCTGGGGAAGAAATGAAGTTATTTGCTCAGACTTTCACATAGAGAGAGTT 360
 DB 301 CGAGTTCCTGGGGAAGAAATGAAGTTATTTGCTCAGACTTTCACATAGAGAGAGTT 360
 QY 361 ACACCTGGCTCCGCCAATCCGCAAAAGTTCCTCATCTCCCTCGGCTCTCTCC 420
 DB 361 ACACCTGGCTCCGCCAATCCGCAAAAGTTCCTCATCTCCCTCGGCTCTCTCC 420
 QY 421 GTTGGCTGGAAGAGATGAGATGCCACCCCGCTCATTAATTAATGATTTATATGCC 480
 DB 421 GTTGGCTGGAAGAGATGAGATGCCACCCCGCTCATTAATTAATGATTTATATGCC 480
 QY 481 ATCTCCAACTGGGCGCAGAGAGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 540
 DB 481 ATCTCCAACTGGGCGCAGAGAGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 540
 QY 541 AGTGTGTGTGCTCAGCTGTGTGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAT 599
 DB 541 AGTGTGTGTGCTCAGCTGTGTGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAT 599

RESULT 2
 BC013551 2198 bp mRNA linear ROD 07-AUG-2002
 LOCUS BC013551
 DEFINITION clone MGC:19348 IMAGE:4236038, mRNA, complete cds.
 ACCESSION BC013551
 VERSION BC013551.1 GI:15488840
 KEYWORDS MGC.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 2198)
 Srausberg, R.
 Direct Submission
 Submitted (04-SEP-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

REMARK
 COMMENT

USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>
 Contact: (Dickson, Mark) mcdepaxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRAK Plate: 25 Row: m Column: 6.
 Location/Qualifiers
 1.2198
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /map="FVB/N"
 /clone="MGC:19348 IMAGE:4236038"
 /issue_type="Kidney, normal. 5 month old male mouse."
 /clone_lib="NCI CGAP_kid14"
 /lab_host="DHI08"
 /note="Vector: PCMV-SPORT6"
 77.673
 /codon_start=1
 /product="Down syndrome critical region homolog 1 (human)"
 /protein_id="AAH13551.1"
 /db_xref="GI:15488841"
 /db_xref="LocustID:54720"
 /translation="MMPRDSYNPSSLIACVANDVSESETRAKESLFRYDKDTT
 FOYKSFKRVINPISADARLRHKEFLGKEMKLYFAQTLLHSGSLAPND
 KQFLISPPASPVGMKQVEDATPVINDLYAISLKGPGKYEHLATDTTPSVVAV
 CESDQENEEEBEEMRKRPKPIORREYPIPHLS"

BASE COUNT 533 a 539 c 569 g 557 t

ORIGIN

Query Match 95.7%; Score 573; DB 10; Length 2198;
 Best Local Similarity 99.7%; Pred. No. 3,98-171;
 Matches 595; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 5 TCCAAAGAACTCCAGCTTGGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 64
 DB 1 TCCAAAGAACTCCAGCTTGGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
 QY 65 CCGGTAAGAAAGCAGATGATTTTAGGAGCTTTAGCTCAATTTTGGCTCCCTGAT 123
 DB 65 CCGGTAAGAAAGCAGATGATTTTAGGAGCTTTAGCTCAATTTTGGCTCCCTGAT 120
 QY 124 TTTGTGTGCAAGAGATGATCTTTCAGCCAAAGTGAAGAGAGAGAGAGAGAGAGAG 183
 DB 124 TTTGTGTGCAAGAGATGATCTTTCAGCCAAAGTGAAGAGAGAGAGAGAGAGAGAG 180
 QY 184 CTTCAGAAATATGACAG 243
 DB 184 CTTCAGAAATATGACAG 240
 QY 244 GATTAACCTTGAGCAACCCCTTATCTGAGCCGATGCCAGCTGCGCTGCACAGACGA 303
 DB 244 GATTAACCTTGAGCAACCCCTTATCTGAGCCGATGCCAGCTGCGCTGCACAGACGA 300
 QY 304 GTTCTGTGGGAAGAAATGAAGTTATTTTGGCTTACATTAAGAGAGAGAGAGAGAG 363
 DB 304 GTTCTGTGGGAAGAAATGAAGTTATTTTGGCTTACATTAAGAGAGAGAGAGAGAG 360
 QY 364 CTGTGCTCGG-CGAATCCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 422
 DB 364 CTGTGCTCGGCGCAATCCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
 QY 423 TGGCTGGAAGAAAGTGAAGATGCCACCCCGCTCATTAATTAATGATTTATATGCCAT 482

```

*|||||
Db 421 TGGCTGGAAACAGTAGAGATGCCACCCCGTCATAAATTACGATCTTTTATATGCCAT 480
Qy 483 CTCAAGCTGGGCGCAGAGAGAAAGTAGTATGAACATGCATGCAGCGACAGACACCACTCCAG 542
Db 481 CTCAAGCTGGGCGCAGAGAGAAAGTAGTATGAACATGCATGCAGCGACAGACACCACTCCAG 540
Qy 543 TGTGCTGTCACCTGTTGTAGAGTGCACCAAGAGAAATGAGAGGAAGAGAGAT 599
Db 541 TGTGCTGTCACCTGTTGTAGAGTGCACCAAGAGAAATGAGAGGAAGAGAGAT 597

RESULT 3
AF263239 2141 bp mRNA linear ROD 30-MAY-2000
LOCUS Mus musculus calcineurin inhibitor mRNA, complete cds,
DEFINITION alternatively spliced.
ACCESSION AF263239
VERSION AF263239.1 GI:8102011
KEYWORDS Mus musculus
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2141)
TITLE Fuentes, J.J., Pritchard, M.A., Pucharcos, C. and Estivill, X.
JOURNAL Down syndrome candidate region 1 (Dscr1), one of three
alternatively spliced exon 1 transcripts
REFERENCE 2 (bases 1 to 2141)
AUTHORS Fuentes, J.J., Pritchard, M.A., Pucharcos, C. and Estivill, X.
TITLE Direct Submission
JOURNAL Submitted (02-MAY-2000) Centre de Genetica Medica i Molecular, IRO,
Avia. Castelldefels Km. 2.7, L'Hospitalet de Llobregat, Barcelona
08907, Spain
FEATURES
source Location/Qualifiers
1..2141
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="16"
31..627
/note="Down syndrome candidate region 1; Dscr1;
alternatively spliced exon 1"
/codon_start=1
/product="calcineurin inhibitor"
/protein_id="AAF72701.1"
/db_xref="GI:8102012"
/translation="MHFRDYSYNFSSLIACVANDVFESETRAKFESLFRDYDKDTT
FOYKFSKRVIRNFNSAADARLRLHKTFLGKEMKLYFAOTLHGSSHLAPPND
KQFLISPPASPPVGMKQVEDATPVINYDLLYAISKLGPGEKYLHAATDTPSVVVHV
CESDQNEEBEEMERMKPKKIOTRRPEYTPHLS"
BASE COUNT 505 a 527 c 558 g 548 t 3 others
ORIGIN
Query Match 87.5%; Score 524.4; DB 10; Length 2141;
Best Local Similarity 99.5%; Pred. No. 1.2e-155;
Matches 547; Conservative 0; Mismatches 1; Indels 2; Gaps 2;
Qy 52 GTTAACGCTGCCCCGTGAAAGAGAGATG-ATTTTAGGACTTTAGCTACAATTTTA 110
Db 2 GTTAACGCTGCCCCGTGAAAGAGAGATGCAATTTTAGGACTATAGCTACAATTTTA 61
Qy 111 GCTCCCTGATGCTGTGTGGCAACAGATGCTCTTCAGCGAAAGTGAGACGAGGCCA 170
Db 62 GCTCCCTGATGCTGTGTGGCAACAGATGCTCTTCAGCGAAAGTGAGACGAGGCCA 121
Qy 171 AATTTGAATCCCTCTTCAGAAATATGACAGCACACCTTCCAGTATTTTAGAGCT 230
Db 122 AATTTGAATCCCTCTTCAGAAATATGACAGCACACCTTCCAGTATTTTAGAGCT 181
Qy 231 TCAACGCTGTCGGATTAATCTTCAGCAACCCCTTATCTGACGCCGATGCCAGCTGCCGC 290
Db 182 TCAACGCTGTCGGATTAATCTTCAGCAACCCCTTATCTGACGCCGATGCCAGCTGCCGC 241

```

```

Qy 291 TGCACAAGACCGAGTTCTCTGGGGAAGGAATGAAGTTGTTATTTTCTCAGACTTTACACA 350
Db 242 TGCACAAGACCGAGTTCTCTGGGGAAGGAATGAAGTTGTTATTTTCTCAGACTTTACACA 301
Qy 351 TAGGAAGTTTACACCTTGGCTCCG-CCAAATCCCGCAAAACAGTTCTCTCATCTCCCTCCGG 409
Db 302 TAGGAAGTTTACACCTTGGCTCCGCCCCAAATCCCGCAAAACAGTTCTCTCATCTCCCTCCGG 361
Qy 410 CCTCTCTCCCTGCTGGCTGGAAACAAGTAGAAGATGCCACCCCGTCATAAAATTAGCATC 469
Db 362 CCTCTCTCCCTGCTGGCTGGAAACAAGTAGAAGATGCCACCCCGTCATAAAATTAGCATC 421
Qy 470 TTTTATATGCCATCTCCAAAGCTGGGCGCAGGAGAGAAGTATGAATGCAATGCAGCGACAG 529
Db 422 TTTTATATGCCATCTCCAAAGCTGGGCGCAGGAGAGAAGTATGAATGCAATGCAGCGACAG 481
Qy 530 ACACCACTCCCAAGTGTGTGTCACAGTGTGTGAGAGTGACCAAGAGAATGAGGAGGAAG 589
Db 482 ACACCACTCCCAAGTGTGTGTCACAGTGTGTGAGAGTGACCAAGAGAATGAGGAGGAAG 541
Qy 590 AGGAAGAGAT 599
Db 542 AGGAAGAGAT 551

RESULT 4
AX365315 597 bp DNA linear PAT 15-FEB-2002
LOCUS Sequence 5 from Patent WO0204491.
DEFINITION AX365315
ACCESSION AX365315
VERSION AX365315.1 GI:18697045
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
TITLE Williams, S.R. and Rothermel, B.
JOURNAL Methods and compositions relating to muscle selective calcineurin
interacting protein (mcip)
Patent: WO 0204491-A 5 17-JAN-2002;
Board of Regents, The University of Texas System (US); Williams,
Sanders R. (US); Rothermel, Beverly (US)
FEATURES
source Location/Qualifiers
1..597
/organism="Mus musculus"
/db_xref="taxon:10090"
1..597
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD23809.1"
/db_xref="GI:18697046"
/translation="MDPRDYSYNFSSLIACVANDVFESETRAKFESLFRDYDKDTT
FOYKFSKRVIRNFNSAADARLRLHKTFLGKEMKLYFAOTLHGSSHLAPPND
KQFLISPPASPPVGMKQVEDATPVINYDLLYAISKLGPGEKYLHAATDTPSVVVHV
CESDQNEEBEEMERMKPKKIOTRRPEYTPHLS"
BASE COUNT 170 a 156 c 142 g 129 t
ORIGIN
Query Match 84.2%; Score 504.4; DB 6; Length 597;
Best Local Similarity 99.6%; Pred. No. 2.2e-149;
Matches 516; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
Qy 83 GATTTAGGACTTTAGCTACAATTTTAGCTCCCTGATTTGCTTGTGGCAACGATGAT 142
Db 4 GATTTAGGACTTTAGCTACAATTTTAGCTCCCTGATTTGCTTGTGGCAACGATGAT 63
Qy 143 GTCTTCAGCGAAAGTGAGACCGAGCGCCAAATTTGAATCCCTCTTCAGAACATATGACAAG 202
Db 64 GTCTTCAGCGAAAGTGAGACCGAGCGCCAAATTTGAATCCCTCTTCAGAACATATGACAAG 123
Qy 203 GACACCACTTCCAGTATTTTAAAGAGCTTCAAAACGTCGCGGATAAACTTCAGCAACCCC 262

```

```

|||||
Db 124 GACACACCTTCCAGTATTTTAAAGCTTCAAAACGTGTCGGATAACTTCAGCAACCCC 183
Oy 263 TTATCTGACGCCGATGCCAGGCTGCGGCTGCACAAGACCCGAGTTCCTGGGGAAGAAATG 322
Db 184 TTATCTGACGCCGATGCCAGGCTGCGGCTGCACAAGACCCGAGTTCCTGGGGAAGAAATG 243
Oy 323 AAGTTGATTTTGGCTCAGACTTTACACATAGGAAGTTCCACCTGGCTCCG-CCAAATCCC 381
Db 244 AAGTTGATTTTGGCTCAGACTTTACACATAGGAAGTTCCACCTGGCTCCG-CCAAATCCC 303
Oy 382 GACAAACAGTTCCTCATCTCCCTCCGCGCTCTCCTCCGTTGGCTGGAAGAAAGTAGAA 441
Db 304 GACAAACAGTTCCTCATCTCCCTCCGCGCTCTCCTCCGTTGGCTGGAAGAAAGTAGAA 363
Oy 442 GATGCCACCCCGCTCATTAATTAAGATCTTTATATGCAATCTCCAAAGCTGGGCGCAGGA 501
Db 364 GATGCCACCCCGCTCATTAATTAAGATCTTTATATGCAATCTCCAAAGCTGGGCGCAGGA 423
Oy 502 GAGAAGTATGAATCTGATGACGAGGACAGACACACCTCCAGTGTGGTCCACGCTGTGT 561
Db 424 GAGAAGTATGAATCTGATGACGAGGACAGACACACCTCCAGTGTGGTCCACGCTGTGT 483
Oy 562 GAGAGTACCAAGAGATGAGAGAGAGAGAGAT 599
Db 484 GAGAGTACCAAGAGATGAGAGAGAGAGAGAT 521

RESULT 5
AF237790 597 bp mRNA linear ROD 12-APR-2000
LOCUS Mus musculus myocyte-enriched calcineurin interactor protein 1
DEFINITION AF237790
ACCESSION AF237790
VERSION AF237790.1 GI:7542528
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 597)
Roethermel,B.A., Vega,R.B., Yang,J., Wu,H., Bassel-Duby,R.S. and
Williams,R.S.
A protein encoded within the Down Syndrome Critical Region is
Enriched in Striated Muscles and Inhibits Calcineurin Signaling
J. Biol. Chem. (2000) in press
2 (bases 1 to 597)
Roethermel,B.A., Vega,R.B., Yang,J., Wu,H., Bassel-Duby,R.S. and
Williams,R.S.
Direct Submision
Submitted (23-PEB-2000) Internal Medicine, University of Texas
Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, TX
75390, USA
FEATURES
source
1..597
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="16"
1..597
/notes="similar to the EST with Genbank Accession Number
AA200984; similar to Homo sapiens DSCR1"
/codon_start=1
/product="myocyte-enriched calcineurin interactor protein
1 splice variant 4"
/protein_id="AA63486.1"
/db_xref="GI:7542528"
/translation="MDFRDFSYNFSLLIACVADVFSSETRAKESLFRYDKDT
FOYKSPKRVINSPNLSADARLRJLHTEFLGKMKLYFAQTLLIGSSLAAPPD
KQRLSPSPAPVGMKQVEDATPVINDLYAISKGPGEKYLHNAOTPTPSVVVH
CESDQNEEESEEMERKRPKRIQVRRREYPIHLS"
CDS
170 a 156 c 142 g 129 c
BASE COUNT
ORIGIN

```

```

Query Match 84.2%; Score 504.4; DB 10; Length 597;
Best Local Similarity 99.6%; Pred. No. 2,2e-149;
Matches 516; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Oy 83 GATTATAGGACCTTACATTAATTTTAACTCCCTGATGCTGTGTGGCAAGAGAT 142
Db 4 GATTATAGGACCTTACATTAATTTTAACTCCCTGATGCTGTGTGGCAAGAGAT 63
Oy 143 GTCTTCAGCGAAGTGAAGACAGGCGCAAAATTGATCCCTTCAGAACATATGACAG 202
Db 64 GTCTTCAGCGAAGTGAAGACAGGCGCAAAATTGATCCCTTCAGAACATATGACAG 123
Oy 203 GACACCACTTCCAGTATTTTAAAGCTTCAACAGTGTCCGATAACTTACGACACCC 262
Db 124 GACACCACTTCCAGTATTTTAAAGCTTCAACAGTGTCCGATAACTTACGACACCC 183
Oy 263 TTATCTGACGCCGATGCCAGGCTGCGGCTGCACAAGACCCGAGTTCCTGGGGAAGAAATG 322
Db 184 TTATCTGACGCCGATGCCAGGCTGCGGCTGCACAAGACCCGAGTTCCTGGGGAAGAAATG 243
Oy 323 AAGTTGATTTTGGCTCAGACTTTACATAGGAAGTTCCACACTGCTCCG-CCAAATCCC 381
Db 244 AAGTTGATTTTGGCTCAGACTTTACATAGGAAGTTCCACACTGCTCCG-CCAAATCCC 303
Oy 382 GACAAACAGTTCCTCATCTCCCTCCGCGCTCTCCTCCGTTGGCTGGAAGAAAGTAGAA 441
Db 304 GACAAACAGTTCCTCATCTCCCTCCGCGCTCTCCTCCGTTGGCTGGAAGAAAGTAGAA 363
Oy 442 GATGCCACCCCGCTCATTAATTAAGATCTTTATATGCAATCTCCAAAGCTGGGCGCAGGA 501
Db 364 GATGCCACCCCGCTCATTAATTAAGATCTTTATATGCAATCTCCAAAGCTGGGCGCAGGA 423
Oy 502 GAGAAGTATGAATCTGATGACGAGGACAGACACCACTCCAGTGTGGTCCACGCTGTGT 561
Db 424 GAGAAGTATGAATCTGATGACGAGGACAGACACCACTCCAGTGTGGTCCACGCTGTGT 483
Oy 562 GAGAGTACCAAGAGATGAGAGAGAGAGAGAT 599
Db 484 GAGAGTACCAAGAGATGAGAGAGAGAGAGAT 521

RESULT 6
CGU60263 2216 bp mRNA linear ROD 03-SEP-1997
LOCUS Cricetulus griseus Adapt78 (adapt78) mRNA, complete cds.
DEFINITION U60263
ACCESSION U60263
VERSION U60263.1 GI:2351390
KEYWORDS
SOURCE Cricetulus griseus.
ORGANISM Cricetulus griseus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
Cricetulus.
1 (bases 1 to 2216)
Crawford,D.R., Leahy,K.P., Abramova,N., Lan,L., Wang,Y. and
Davies,K.J.
Hamster adapt78 mRNA is a Down syndrome critical region homologue
that is inducible by oxidative stress
Arch. Biochem. Biophys. 342 (1), 6-12 (1997)
JOURNAL 97329095
MEDLINE 9185608
PUBMED 2 (bases 1 to 2216)
REFERENCE Crawford,D.R., Leahy,K.L. and Davies,K.J.A.
AUTHORS Direct Submision
Submitted (10-JUN-1996) Biochem. and Mol. Biol., Albany Medical
College, 47 New Scotland Avenue, Albany, NY 12208, USA
FEATURES
source
1..2216
/organism="Cricetulus griseus"
/db_xref="taxon:10029"
/cell_type="HA-1"
1..2216
/gene="adapt78"

```

CDS	70..663	/gene="adapt78"	
		/note="similar to human DSCR1; an oxidant- and calcium-inducible mRNA"	
		/codon_start=1	
		/product="Adapt78"	
		/protein_id="AA068517.1"	
		/db_xref="GI:2351391"	
		/translation="MHFRDFNYFSSLIACVANGDVSESETRAKFESLFRVYDKDIT FQYFKSPKRVIRINFSNPLSAADARLQHLKTEFLGKEMKLYFAOTLHGSSHLAPPND KQFLISPPAPPVGMKQVEDATPVINDLYAISKLGPGEKYLHAATDTTSPVVHVH CESDQNEEEBEEEMERMKRPKIKIOTRRPEYTPHLS"	
BASE COUNT	544 a 515 c 564 g 593 t		
ORIGIN			
Query Match	75.7%; Score 453.2; DB 10; Length 2216;		
Best Local Similarity	90.2%; Pred. No. 6.4e-133;		
Matches	507; Conservative 0; Mismatches 53; Indels 2; Gaps 2;		
Qy	34	TGAGAGAGCGAGTCGTTGCTTAAGCGCTGCCCCGCTGAAAAGCAGAATG-ATTTTAGGG 92	
Db	23	TGAGAGAGAGTCGTGAGTCATTAAGCGCTGCGCTGTGAAGAAAGCAGAATGCAATTTTAGGG 82	
Qy	93	ACTTTAGCTACAAATTTTAGCTCCCTGATGCTGTGTGGCAACGATGATGCTTCAGCG 152	
Db	83	ACTTTAACTACAAATTTTAGCTCCCTGATGCTGTGTGGCAACGATGATGCTTCAGCG 142	
Qy	153	AAAGTGAGACGAGCGCCAAATTTGAATCCCTCTTCAGAACATATGACAGCACCAACCT 212	
Db	143	AAAGTGAAACGAGCGCCAAATTTGAATCCCTCTTCAGAGCGTATGACAGGACATCACTT 202	
Qy	213	TCCAGTATTTTAAAGAGCTTCAAAACGCTCCCGATTAACCTTCAGCAACCCCTTATCTGCAG 272	
Db	203	TCCAGTATTTTAAAGAGCTTCAAAACGCTTCCGAAATAACTTCAGCAACCCCTTATCCGCAG 262	
Qy	273	CCGATGCCAGGCTCGCGCTGCACAAGACCGAGTTCTCTGGGGAAGGAAATGAAGTCTGATT 332	
Db	263	CTGACGCGAGGCTGCAGCTGCATTAAGACGAGTTCCTCTGGGGAAGGAAATGAAGTCTGATT 322	
Qy	333	TTGCTCAGAGCTTTACACATAGGAAGTTACACCTGGCTCCG-CCAAATCCGACAAACAGT 391	
Db	323	TTGCTCAGAGCTTTACACATAGGAAGTTACACCTGGCTCCGCGCCATCCAGACAGCACT 382	
Qy	392	TCCTCATCTCCCTCCGCGCTCTCTCCCGCTTGGCTGGGAAACAAGTAGAAGATGCCACCC 451	
Db	383	TCCTCATCTCTCTCCGCGCTCTCCACAGTGGTTGGGAAGCAAGTAGAAGACGCTACCC 442	
Qy	452	CCGTCATAAATTACGATCTTTTATATGCGATCTTCAAGCTGGGGCCAGGAGAGAAAGTATG 511	
Db	443	CAGTCATAAATTACGATCTTTTATATGCTATCTTCAAGCTGGGGCCAGGTGAGAAAGTATG 502	
Qy	512	AACCTGATCGAGCGACAGACACCACTCCAGTGTGGTGTGTCACAGTGTGTGAGAGTGACC 571	
Db	503	AGCTGATCGAGCAACAGACACCACTCCAGTGTGGTGTGTCACAGTGTGTGAGAGTGACC 562	
Qy	572	AAGAGAAATCAGGAGGAGGAGGA 593	
Db	563	AAGAGAAATCAGGAGGAGGAGGA 584	
RESULT 7			
AB075973			
LOCUS	AB075973	626 bp mRNA linear	ROD 27-JUL-2002
DEFINITION	Rattus norvegicus mcip 1 mRNA for myocyte-enriched calcineurin-interacting protein 1, complete cds.		
ACCESSION	AB075973		
VERSION	AB075973.1	GI:21998843	
KEYWORDS	Rattus norvegicus cdna to mRNA.		
SOURCE	Rattus norvegicus		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
REFERENCE	1		
AUTHORS	Odashima,M., Nagata,K., Obata,K., Somura,F., Izawa,H., Miyazaki,T., Murata,Y., Seo,H., Yamada,Y. and Yokota,M.		
TITLE	rat myocyte-enriched calcineurin interacting protein 1, splice variant 4 mRNA, complete cds		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 626)		
AUTHORS	Obata,K., Nagata,K., Odashima,M., Somura,F., Yamada,Y. and Yokota,M.		
TITLE	Direct Submission		
JOURNAL	Submitted (11-DEC-2001) Koji Obata, Nagoya University, Department of Clinical Pathophysiology; 65 Teurumai-cho, Showa-ku, Nagoya, Aichi 466-8550, Japan (E-mail:obata@etsu.med.nagoya-u.ac.jp, Tel:81-52-744-2577, Fax:81-52-744-2977)		
FEATURES	Location/Qualifiers		
source	1..626		
gene	/organism="Rattus norvegicus"		
	/db_xref="taxon:10116"		
CDS	1..626		
	/gene="mcip 1"		
	19..618		
	/gene="mcip 1"		
	/codon_start=1		
	/product="myocyte-enriched calcineurin-interacting protein 1"		
	/protein_id="BAC06443.1"		
	/db_xref="GI:21998844"		
	/translation="MHFRDFNYFSSLIACVANGDVSESETRAKFESLFRVYDKDIT FQYFKSPKRVIRINFSNPLSAADARLQHLKTEFLGKEMKLYFAOTLHGSSHLAPPND KQFLISPPAPPVGMKQVEDATPVINDLYAISKLGPGEKYLHAATDTTSPVVHVH CESDQNEEEBEEEMERMKRPKIKIOTRRPEYTPHLS"		
BASE COUNT	182 a 163 c 154 g 127 t		
ORIGIN			
Query Match	72.8%; Score 435.8; DB 10; Length 626;		
Best Local Similarity	91.7%; Pred. No. 1.7e-127;		
Matches	472; Conservative 0; Mismatches 42; Indels 1; Gaps 1;		
Qy	84	ATTTTAGGAGCTTTAGCTACAAATTTTAGCTCCCTGATGCTTGTGTGGCAACGATGATG 143	
Db	23	ATTTTAGGAGCTTTAACTACAAATTTTAGCTCCCTGATGCTTGTGTGGCAACGATGATG 82	
Qy	144	TCCTCAGCGAAGTGAGACCAAGGCGCCAAATTTGAATCCCTCTTCAGAACATATGACAGG 203	
Db	83	TCCTCAGCGAAGTGAGACCAAGGCGCCAAATTTGAATCCCTCTTCAGAACATATGACAGG 142	
Qy	204	ACACACCTTCCAGTATTTTAAGAGCTTCAAAAGTGTCCGATATAAATTCAGCAACCCCT 263	
Db	143	ACATCACTTCCAGTATTTTAAGAGCTTCAAGCGGTGTCGAATAAACTTCAGCAACCCCT 202	
Qy	264	TATCTGAGCGCGATGCCAGGCTGCGGCTGCACAGACCGAGTTCCTGGGGAAGAAATGA 323	
Db	203	TGCTGGGCGGACCGCAGGCTGCGGCTGCATAGACAGAGTTCCTGGGGAAGAGATGA 262	
Qy	324	AGTTGTATTTTGTCTCAGACTTTACACATAGGAAGTTACACCTGGCTCCG-CCAAATCCCG 382	
Db	263	AACGTACTTTTGCCAGACTTTACACATAGGAAGTTACACCTGGCTCCGCGCCAAATCCAG 322	
Qy	383	ACAAACAGTTTCTCATCTCCCTCCGCGCTCTCTCCCGTGGCTGGAACAAAGTAGAAG 442	
Db	323	ACAAACAGTTTCTCATCTCTCCCGCTCTCCAGCTGTGGCTGGAACAAAGTAGAAG 382	
Qy	443	ATGCCACCCCGCTCATAAATTTAGCATCTTTATATGCCATCTCCAGCTGGGCGCAGGAG 502	
Db	383	ACGCTACCCCGCTCATAAATTTAGCATCTTTATATGCTATCTCCAAAGCTGGGACCGAGG 442	
Qy	503	AGAAGTATGAATGCGATGCGGACAGACACCACTCCAGTGTGTGTGGTCCAGCTGTGTG 562	
Db	443	AGAAGTATGAGCTACACGAGGACAGACACCAAGCTGTGTGTGGTCCAGCTGTGTG 502	
Qy	563	AGAGTCAACCAAGAGATGAGGAGGAGAGAGAG 597	
Db	503	AGAGTCAACCAAGAGGAGGAGGAGAGAGAGAG 537	

RESULT 8
AF282255 621 bp mRNA linear ROD 17-NOV-2000
LOCUS AF282255
DEFINITION Mus musculus Down syndrome candidate region 1 protein (Dscr1) mRNA,
complete cds.
ACCESSION AF282255
VERSION AF282255.1 GI:9652249
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 621)
Stripoli, P., Petrini, M., Lenzi, L., Carinci, P. and Zannotti, M.
The murine DSCR1-like (Down syndrome candidate region 1) gene
family: conserved syteny with the human orthologous genes
JOURNAL Gene 257 (2), 223-232 (2000)
MEDLINE 20534792
PUBMED 11080588
REFERENCE 2 (bases 1 to 621)
Stripoli, P., Petrini, M., Lenzi, L., Carinci, P. and Zannotti, M.
Direct Submision
Submitted (24-JUN-2000) Istituto di Istologia ed Embriologia
Generale, Universita di Bologna, Via Belmeloro, 8, Bologna, BO
40126, Italy

FEATURES

Source

1..621
/organism="Mus musculus"
/strain="BALB/c"
/db_xref="taxon:10090"
/chromosome="16"
/sex="male"
/cruise_type="brain"
1..621
/gene="Dscr1"
8..604
/gene="Dscr1"
/note="similar to the Homo sapiens Down syndrome candidate
region 1 protein"
/codon_start=1
/product="Down syndrome candidate region 1 protein"
/protein_id="AA91461.1"
/db_xref="GI:9652250"
/translation="MEVDLDQPLPSATICHLDPRFVVDGLCAKRESLFRTYDKDT
FOYKSPKRVINFSNPLSADARLRLHKTPELGKEMKLYPAOTLHIGSHLAPVPD
KQPLISPPSPVGMKQVEDATPVINDLVLSKLGPEKEELHAAITPTTSVVHV
CESDQNEEBEEMERMRPKRIOTRREYTPHLS"

gene

CDS

BASE COUNT 166 a 179 c 159 g 117 t
ORIGIN

Query Match 71.4%; Score 427.8; DB 10; Length 621;
Best Local Similarity 93.3%; Pred. No. 6.1e-125;
Matches 458; Conservative 0; Mismatches 32; Indels 1; Gaps 1;

QY 110 AGCTCCCTGATTGTTGTGTGCAACGATGATCTTTCAGGGAAGTGAAGCAGGCGC 169
DB 38 AGCGCACATGCGCTGCGACCTGACCGCGCGTGTGTGTGAGCGGCTGCGGAGC 97
QY 170 AATTGTAATCCCTTTCAGAACATATGACAGACACCACTTCAGATATTTAAGAGC 229
DB 98 AATTGTAATCCCTTTCAGAACATATGACAGACACCACTTCAGATATTTAAGAGC 157
QY 230 TTCAACGTCGCGATTAATCTTACGCAACCCCTTATCTGACCGGATGCCAGCTGGG 289
DB 158 TTCAACGTCGCGATTAATCTTACGCAACCCCTTATCTGACCGGATGCCAGCTGGG 217
QY 290 CTGACACAGACCGAGTTCTCTGGGGAAGAAATGAAGTTTATTTTGGCTGAGACTTACAC 349
DB 218 CTGACACAGACCGAGTTCTCTGGGGAAGAAATGAAGTTTATTTTGGCTGAGACTTACAC 277
QY 350 ATTAGAAGTTTACACCTGGCTCG-CCATCCCGCAACAGTCTCTCCCTCGC 408

DB 278 ATGGAAGTTACACCTCGGCTCGCCCAATCCGACAAACAGTTCCTCATCTCCCTCGC 337
QY 409 GCCTCTCTCTCCGTTGGCTGGGAACAAGTAGATGCCACCCGTCATTAATACGAT 468
DB 338 GCCTCTCTCTCCGTTGGCTGGGAACAAGTAGATGCCACCCGTCATTAATACGAT 397
QY 469 CTTTATATGTCATCTCCAGAGCTGGGGCCAGAGAGAGTATGACTGCATGACGCGACA 528
DB 398 CTTTATATGTCATCTCCAGAGCTGGGGCCAGAGAGAGTATGACTGCATGACGCGACA 457
QY 529 GACACCACTCCCACTGTCGTGTGTCACGTGTGAGAGTACCAAGAAATGAGAGGAA 588
DB 458 GACACCACTCCCACTGTCGTGTGTCACGTGTGAGAGTACCAAGAAATGAGAGGAA 517
QY 589 GAGGAAGAGAT 599
DB 518 GAGGAAGAGAT 528

RESULT 9
AF260717 2125 bp mRNA linear ROD 22-MAR-2001
LOCUS AF260717
DEFINITION Mus musculus Down syndrome candidate region 1 (Dscr1) mRNA, complete
cds.
ACCESSION AF260717
VERSION AF260717.1 GI:7839596
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2125)
Canaas, C., Martinez, S., Pritchard, M.A., Fuentes, J.J., Nadal, M.,
Guilera, J., Arbones, M., Flores, J., Soriano, E., Bativilla, X. and
Alcantara, S.
Dscr1, a novel endogenous inhibitor of calcineurin signaling, is
expressed in the primitive ventricle of the heart and during
neurogenesis
Mol. Cell. Dev. 101 (1-2), 289-292 (2001)

FEATURES

Source

1..2125
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="16"
1..2125
/gene="Dscr1"
15..611
/gene="Dscr1"
/note="calcineurin inhibitor"
/codon_start=1
/product="Down syndrome candidate region 1"
/protein_id="AA91461.1"
/db_xref="GI:7839597"
/translation="MEVDLDQPLPSATICHLDPRFVVDGLCAKRESLFRTYDKDT
FOYKSPKRVINFSNPLSADARLRLHKTPELGKEMKLYPAOTLHIGSHLAPVPD
KQPLISPPSPVGMKQVEDATPVINDLVLSKLGPEKEELHAAITPTTSVVHV
CESDQNEEBEEMERMRPKRIOTRREYTPHLS"

gene

CDS

BASE COUNT 487 a 536 c 563 g 536 t 3 others
ORIGIN

Query Match 71.4%; Score 427.8; DB 10; Length 621;
Best Local Similarity 93.3%; Pred. No. 6.1e-125;
Matches 458; Conservative 0; Mismatches 32; Indels 1; Gaps 1;

QY 110 AGCTCCCTGATTGTTGTGTGCAACGATGATCTTTCAGGGAAGTGAAGCAGGCGC 169
DB 38 AGCGCACATGCGCTGCGACCTGACCGCGCGTGTGTGTGAGCGGCTGCGGAGC 97
QY 170 AATTGTAATCCCTTTCAGAACATATGACAGACACCACTTCAGATATTTAAGAGC 229
DB 98 AATTGTAATCCCTTTCAGAACATATGACAGACACCACTTCAGATATTTAAGAGC 157
QY 230 TTCAACGTCGCGATTAATCTTACGCAACCCCTTATCTGACCGGATGCCAGCTGGG 289
DB 158 TTCAACGTCGCGATTAATCTTACGCAACCCCTTATCTGACCGGATGCCAGCTGGG 217
QY 290 CTGACACAGACCGAGTTCTCTGGGGAAGAAATGAAGTTTATTTTGGCTGAGACTTACAC 349
DB 218 CTGACACAGACCGAGTTCTCTGGGGAAGAAATGAAGTTTATTTTGGCTGAGACTTACAC 277
QY 350 ATTAGAAGTTTACACCTGGCTCG-CCATCCCGCAACAGTCTCTCCCTCGC 408


```
Query Match      71.4%; Score 427.8; DB 10; Length 2125;
Best Local Similarity 93.3%; Pred. No. 8e-125;
Matches 458; Conservative 0; Mismatches 32; Indels 1; Gaps 1;

QY 110 AGCTCCCTGATTCCTGTCGCAACGATGATGCTTCAGCGAAAGTGAGACGAGGCC 169
DB 45 AGCGCCACCATCGCTGCGCACCTGACCGCGCTGTCGTGAGCGGCTGTGCGCGGCC 104
QY 170 AAATTGTAATCCCTCTTCAGAACATATGACAGGACACACCTTCAGTATTTAAGAGC 229
DB 105 AAATTGTAATCCCTCTTCAGAACATATGACAGGACACACCTTCAGTATTTAAGAGC 164
QY 230 TTCAACGTCGCGATTAACCTTCAGCAACCCCTTATCTGACGCGGATGCCAGGCTGCGG 289
DB 165 TTCAACGTCGCGATTAACCTTCAGCAACCCCTTATCTGACGCGGATGCCAGGCTGCGG 224
QY 290 CTGCAACAGACCGAGTTCCTGGGGAAGAAATGAGTTGTTATTTGCTCAGACTTTACAC 349
DB 225 CTGCAACAGACCGAGTTCCTGGGGAAGAAATGAGTTGTTATTTGCTCAGACTTTACAC 284
QY 350 ATAGGAAGTTCACACTGGCTCCG - CCAATCCGACAAACAGTTCCTCATCTCCCTCCG 408
DB 285 ATAGGAAGTTCACACTGGCTCCG - CCAATCCGACAAACAGTTCCTCATCTCCCTCCG 344
QY 409 GCCTCTCTCCGTTGGCTGGAACAGTGAAGATGCCACCCCGTCATAAATTACGAT 468
DB 345 GCCTCTCTCCGTTGGCTGGAACAGTGAAGATGCCACCCCGTCATAAATTACGAT 404
QY 469 CTTTATATGCACTTCCAACTGGGCGGAGAGATGATGAAGTTCATGTCAGCGACA 528
DB 405 CTTTATATGCACTTCCAACTGGGCGGAGAGATGATGAAGTTCATGTCAGCGACA 464
QY 529 GACACCACTCCAGTGGTGGTCCAGTGTGTGAGAGTGACCAAGAGATGAGGAGAA 588
DB 465 GACACCACTCCAGTGGTGGTCCAGTGTGTGAGAGTGACCAAGAGATGAGGAGAA 524
QY 589 GAGGAAGAGAT 599
DB 525 GAGGAAGAGAT 535

RESULT 10
AX365312
LOCUS      AX365312
DEFINITION Sequence 2 from Patent WO0204491.
ACCESSION AX365312
VERSION    AX365312.1 GI:18697043
KEYWORDS   house mouse.
SOURCE     Mus musculus
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS    Williams, S.R. and Rothermel, B.
TITLE      Methods and compositions relating to muscle selective calcineurin
           interacting protein (mcip)
JOURNAL    Patent: WO 0204491-A 2.17-JAN-2002;
           Board of Regents, The University of Texas System (US); Williams,
           Sanders R. (US); Rothermel, Beverly (US)
FEATURES
source     Location/Qualifiers
           1..597
           /organism="Mus musculus"
           /db_xref="taxon:10090"
           /note="unnamed protein product"
           /codon_start=1
           /protein_id="CAD23808.1"
           /db_xref="GI:18697044"
           /translation="NEEVDLQDLSATIACHLDRFVDFGLCRAPFSLRFTYDKDTT
           FOYKFSFRVRINFSNPUSAADARLRLHKTFLGKEMKLYPAQTLHIGSSHLAPPND
           KQFLISPPASPPVGNKQVEDATPVINTDLYAISKLGPGKEYELHAATDPTPSVVVHV
           CESDQNEEEEEEMRKRPKPIQTRRPEVTPILHS"

BASE COUNT      160 a      170 c      151 g      116 t
ORIGIN
Query Match      71.2%; Score 426.2; DB 6; Length 597;
Best Local Similarity 93.1%; Pred. No. 2e-124;
Matches 457; Conservative 0; Mismatches 33; Indels 1; Gaps 1;

QY 110 AGCTCCCTGATTCCTGTCGCAACGATGATGCTTCAGCGAAAGTGAGACGAGGCC 169
DB 31 AGCGCCACCATCGCTGCGCACCTGACCGCGCTGTCGTGAGCGGCTGTGCGCGGCC 90
QY 170 AAATTGTAATCCCTCTTCAGAACATATGACAGGACACACCTTCAGTATTTAAGAGC 229
DB 91 AAATTGTAATCCCTCTTCAGAACATATGACAGGACACACCTTCAGTATTTAAGAGC 150
QY 230 TTCAACGTCGCGATTAACCTTCAGCAACCCCTTATCTGACGCGGATGCCAGGCTGCGG 289
DB 151 TTCAACGTCGCGATTAACCTTCAGCAACCCCTTATCTGACGCGGATGCCAGGCTGCGG 210
QY 290 CTGCAACAGACCGAGTTCCTGGGGAAGAAATGAGTTGTTATTTGCTCAGACTTTACAC 349
DB 211 CTGCAACAGACCGAGTTCCTGGGGAAGAAATGAGTTGTTATTTGCTCAGACTTTACAC 270
QY 350 ATAGGAAGTTCACACTGGCTCCG - CCAATCCGCAACAGTTCCTCATCTCCCTCCG 408
DB 271 ATAGGAAGTTCACACTGGCTCCG - CCAATCCGCAACAGTTCCTCATCTCCCTCCG 330
QY 409 GCCTCTCTCCGTTGGCTGGAACAGTGAAGATGCCACCCCGTCATAAATTACGAT 468
DB 331 GCCTCTCTCCGTTGGCTGGAACAGTGAAGATGCCACCCCGTCATAAATTACGAT 390
QY 469 CTTTATATGCACTTCCAACTGGGCGGAGAGATGATGAAGTTCATGTCAGCGACA 528
DB 391 CTTTATATGCACTTCCAACTGGGCGGAGAGATGATGAAGTTCATGTCAGCGACA 450
QY 529 GACACCACTCCAGTGGTGGTCCAGTGTGTGAGAGTGACCAAGAGATGAGGAGAA 588
DB 451 GACCCCACTCCAGTGGTGGTCCAGTGTGTGAGAGTGACCAAGAGATGAGGAGAA 510
QY 589 GAGGAAGAGAT 599
DB 511 GAGGAAGAGAT 521
```

RESULT 11

```
AX365312
LOCUS      AF237789
DEFINITION Mus musculus myocyte-enriched calcineurin interactin protein 1
ACCESSION AF237789
VERSION    AF237789.1 GI:7542525
KEYWORDS   Mus musculus.
SOURCE     Mus musculus
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS    Rothermel, B.A., Vega, R.B., Yang, J., Wu, H., Bassel-Duby, R.S. and
           Williams, R.S.
TITLE      A Protein Encoded within the Down Syndrome Critical Region is
           Enriched in Striated Muscles and Inhibits Calcineurin Signaling
JOURNAL    J. Biol. Chem. (2000) In press
REFERENCE   2 (bases 1 to 597)
AUTHORS    Rothermel, B.A., Vega, R.B., Yang, J., Wu, H., Bassel-Duby, R.S. and
           Williams, R.S.
TITLE      Direct Submission
JOURNAL    Submitted (23-FEB-2000) Internal Medicine, University of Texas
           Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, TX
           75390, USA
FEATURES
source     Location/Qualifiers
           1..597
           /organism="Mus musculus"
           /db_xref="taxon:10090"
```

CDS

/chromosome="16"
 1..597
 /note="similar to Homo sapiens DSCR1; NCIP1; similar to
 EST sequence of GenBank Accession Number AA734360"
 /codon_start=1
 /product="myocyte-enriched calcineurin interactin protein
 1 splice variant 1"
 /protein_id="AA63485.1"
 /db_xref="GI:7542526"
 /translation="MEVDLQDLPSPATIAHLDRVVDGLCPAKESLFRTYDKOTT
 FOYKSPKRVYINFSPLAADARLRLKTEFLGKEMKLYFAQPLHSGSLAPNPD
 KQPLSPSPVPGVKQVEDATPVYNLDLYAISKLGPGEYELHAATDPTPSVVHV
 CESDQENBEEBEMERKRPKIKIQRREPYTTHLS"

BASE COUNT

160 a 170 c 151 g 116 t

ORIGIN

Query Match 71.2%; Score 426.2; DB 10; Length 597;
 Best Local Similarity 93.1%; Pred. No. 2e-124;
 Matches 457; Conservative 0; Mismatches 33; Indels 1; Gaps 1;

OY 110 AGCTCCCTGATGCTGTGTGGCAACGATGATGCTTTCAGGAAAGTGACCGGGCC 169
 DB 31 AGCGCACCATGCGCTGCGACCTGACCGCGCGCTGTCTGTGACGCGCTGTGCGGGCC 90
 OY 170 AATTGGAATCCCTCTTCAGAACATATGACAGACACCTCCAGTATTTTAAGAGC 229
 DB 91 AATTGGAATCCCTCTTCAGAACATATGACAGACACCTCCAGTATTTTAAGAGC 150
 OY 230 TTCAAACTGTCCGCGATAAATTGACGAAACCCCTTATCTGACGCGATGCCAGGCTGCGG 289
 DB 151 TTCAAACTGTCCGCGATAAATTGACGAAACCCCTTATCTGACGCGATGCCAGGCTGCGG 210
 OY 290 CTGACACAGACCGAGTTCTGTGGGAAAGAAATGAAATTTTATTTTGTCTAGACTTTAC 349
 DB 211 CTGACACAGACCGAGTTCTGTGGGAAAGAAATGAAATTTTATTTTGTCTAGACTTTAC 270
 OY 350 ATAGGAAGTTCAACCTGGCTCCG-CCAAATCCCGCAACAGTTCTCTATCTCCCTCCG 408
 DB 271 ATAGGAAGTTCAACCTGGCTCCG-CCAAATCCCGCAACAGTTCTCTATCTCCCTCCG 330
 OY 409 GCCTCTCTCCCTGTTGGCTGGAAACAAGTAGAAGATGCCACCCCGCTCATTAATTACAT 468
 DB 331 GCCTCTCTCCCTGTTGGCTGGAAACAAGTAGAAGATGCCACCCCGCTCATTAATTACAT 390
 OY 469 CTTTATATGCCATCTTCGAAGCTGGGGCCGAGAGAGAAAGTATGAATGATGAGAGACA 528
 DB 391 CTTTATATGCCATCTTCGAAGCTGGGGCCGAGAGAGAAAGTATGAATGATGAGAGACA 450
 OY 529 GACACCACTCCAGTGTGTGTGTCACAGTGTGAGAGTGAACCAAGAGATGAGAGAGA 588
 DB 451 GACACCACTCCAGTGTGTGTGTCACAGTGTGAGAGTGAACCAAGAGATGAGAGAGA 510
 OY 589 GAGAGAGAGAT 599
 DB 511 GAGAGAGAGAT 521

RESULT 12
 AF263240 2224 bp mRNA linear ROD 30-MAY-2000
 LOCUS AF263240
 DEFINITION Mus musculus calcineurin inhibitor mRNA, complete cds,
 alternatively spliced.
 ACCESSION AF263240.1 GI:8102013
 VERSION AF263240.1
 KEYWORDS Mus musculus
 SOURCE Mus musculus
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 2224)
 Reference: Puentes, J.J., Pritchard, M.A., Pucharcos, C. and Estivill, X.
 Down syndrome candidate region 1 (Dscr1), one of three
 alternatively spliced exon 1 transcripts

JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 2224)
 AUTHORS Puentes, J.J., Pritchard, M.A., Pucharcos, C. and Estivill, X.
 TITLE Direct Submission
 JOURNAL Submitted (02-MAY-2000) Centre de Genetica Medica i Molecular, IRO,
 Avia, Castelldefels Km. 2,7, L'Hospitalat de Llobregat, Barcelona
 08907, Spain

FEATURES

Location/Qualifiers
 source
 1..2224
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /chromosome="16"
 354..710
 /note="Down syndrome candidate region 1; Dscr1;
 alternatively spliced exon 1"
 /codon_start=1
 /product="calcineurin inhibitor"
 /protein_id="AA63485.1"
 /db_xref="GI:8102014"
 /translation="MKLYFAQPLHSGSLAPNPDKQPLSPSPVPGVKQVEDAT
 PVYNLDLYAISKLGPGEYELHAATDPTPSVVHVCSDDQENBEEBEMERKRPK
 KIKIQRREPYTTHLS"

CDS

BASE COUNT 519 a 554 c 586 g 562 t 3 others
 ORIGIN
 Query Match 71.0%; Score 425.4; DB 10; Length 2224;
 Best Local Similarity 99.5%; Pred. No. 4.7e-124;
 Matches 437; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

OY 162 CCAGGGCCAAATTTGAATCCCTCTTCAGAACATATGACAGACACCACTTCAGATTT 221
 DB 196 CCAGGGCCAAATTTGAATCCCTCTTCAGAACATATGACAGACACCACTTCAGATTT 255
 OY 222 TTAAGAGCTTCAACGTCGCGATTAACCTTCAGCAACCCCTTATCTGACGCGATGCCA 281
 DB 256 TTAAGAGCTTCAACGTCGCGATTAACCTTCAGCAACCCCTTATCTGACGCGATGCCA 315
 OY 282 GGTGCGGCTGCACAGACCGAGTTCTGTGGGAAAGAAATGAAATTTTATTTTGTCTAGA 341
 DB 316 GGTGCGGCTGCACAGACCGAGTTCTGTGGGAAAGAAATGAAATTTTATTTTGTCTAGA 375
 OY 342 CTTTACATATGAAAGTTCAACCTGGCTCCG-CCAAATCCCGCAACAGTTCTCTCATCT 400
 DB 376 CTTTACATATGAAAGTTCAACCTGGCTCCG-CCAAATCCCGCAACAGTTCTCTCATCT 435
 OY 401 CCCCTCCGCGCTCTCTCCCTGCTGGTGAACAAGTAGAAGATGCCACCCCGCTCATTA 460
 DB 436 CCCCTCCGCGCTCTCTCCCTGCTGGTGAACAAGTAGAAGATGCCACCCCGCTCATTA 495
 OY 461 ATTACGATCTTTTATATGCCATCTTCGAAGCTGGGGCCGAGAGAGAAAGTATGAATGATG 520
 DB 496 ATTACGATCTTTTATATGCCATCTTCGAAGCTGGGGCCGAGAGAGAAAGTATGAATGATG 555
 OY 521 CAGCGACAGACCACTCCAGTGTGTGTGTCACAGTGTGTGAGAGTGAACCAAGATG 580
 DB 556 CAGCGACAGACCACTCCAGTGTGTGTGTCACAGTGTGTGAGAGTGAACCAAGATG 615
 OY 581 AGGAGAGAGAGAGAGAT 599
 DB 616 AGGAGAGAGAGAGAGAT 634

RESULT 13

AX365321 2331 bp DNA linear PAT 15-FEB-2002
 LOCUS AX365321
 DEFINITION Sequence 11 from Patent WO2004491.
 ACCESSION AX365321
 VERSION AX365321.1 GI:18697049
 KEYWORDS human.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

1
REFERENCE*
AUTHORS Williams,S.R. and Rothermel,B.
TITLE Methods and compositions relating to muscle selective calcineurin
interacting protein (mcip)
JOURNAL Patent: WO 0204491-A 11 17-JAN-2002;
Board of Regents, The University of Texas System (US) ; Williams,
Sanders R. (US) ; Rothermel, Beverly (US)
FEATURES Location/Qualifiers
source 1..2331
/organism="Homo sapiens"
/db_xref="taxon:9606"
144..737
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD23811.1"
/translation="MHFRNFYSFSLIACVANSDFSESETRAKESLPRTYDKDIT
FOYFKSPKRVIRNFSPFSAADARLQHLKTEFLGKMKLYFACTLHIGSSHLAPPND
KQFLISPPASPPVGMKQVEDATPVINYDLIYAIKSLGPKGYELHAATDTTSPVVHV
CESDQKEEEEEEMRRPKKIIQTRRPEYTHLS"
BASE COUNT 630 a 470 c 547 g 684 t
ORIGIN
Query Match 69.1%; Score 413.8; DB 6; Length 2331;
Best Local Similarity 85.1%; Pred. No. 2.4e-120;
Matches 474; Conservative 0; Mismatches 82; Indels 1; Gaps 1;
Qy 39 GAGCAGTGTTCCTTAAGCGTCTGCCCGTGAAAAAGCAGAATGATTTTAGGACTTTA 98
Db 103 GACTCGTGGTCTGTAGCGCTTTCACGTGAAGAAAGCAAGATGCAATTTAGAACTTTA 162
Qy 99 GCTACAAATTTAGCTCCCTGATTCCTTGTGTGGCAACAGATGATGCTTCAGCGAAAGTG 158
Db 163 ACTACAGTTTTCCTGCTGATTCCTTGTGTGGCAACAGATGATGCTTCAGCGAAAGTG 222
Qy 159 AGACGAGGGCCAAATTTGAATTCCTCTTCAGAACATATGACAGGACACCACTTCCAGT 218
Db 223 AAACGAGGGCCAAATTTGAGTCCCTCTTTAGGACGTATGACAGGACATCACCTTTTCAGT 282
Qy 219 ATTTTAGAGCTTCAACAGTGTCCGGATAAATTCAGCAACCCCTTATCTGACGCGATG 278
Db 283 ATTTTAGAGCTTCAACAGTGTCCGGATAAATTCAGCAACCCCTTATCTGACGCGATG 342
Qy 279 CCAGGCTCGCGCTGCACAGCAGCGAGTCTCTGGGAGGAGGAATGAAGTGTATTTGCTC 338
Db 343 CCAGGCTCGAGCTGCATAGACGTAGTTCTCGGAAGGAAATGAAGTGTATTTGCTC 402
Qy 339 AGACTTTACATAGGAAGTTACACCTGGCTCCGCC-AAATCCGCAACACAGTTCCTCA 397
Db 403 AGACTTTACATAGGAAGTTCACACCTGGCTCCGCCAAATCCAGACAGCAGTTTCTCA 462
Qy 398 TCTCCCTCCGGCTCTCCTCCGTTGGCTGGAAACAAGTAGAAGATGCCACCCCGTCA 457
Db 463 TCTCCCTCCGGCTCTCCTCCGCTGGAGTGGAAACAAGTGGAAAGATGCCACCCCGTCA 522
Qy 458 TAAATACGATCTTTTATATGTCATCTCCAGCTGGGGCCAGAGAGATGATGACTGC 517
Db 523 TAAATATGATCTTTATATGTCATCTCCAGCTGGGGCCAGAGAGATGATGAAATTCG 582
Qy 518 ATCAGCGGACAGACACATCTCCAGTGTGGTGTCCAGCTGTGTGAGAGTGACCAAGAGA 577
Db 583 ACAGCGGACTGACACCACTCCCGAGGTGGTGTCTCATGTATGTGAGATGATCAAGAGA 642
Qy 578 ATGAGGAGGAAGAGAA 594
Db 643 AGAGGAAGGAAGAGAA 659
RESULT 14
HSU85267
LOCUS
DEFINITION Homo sapiens down syndrome candidate region 1 (DSCR1) gene,
alternative exon 1, complete cds.
2346 bp mRNA linear PRI 19-APR-2000

U85267
U85267.2 GI:7596915
Homo sapiens
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2346)
Fuentes,J.J., Pritchard,M.A., Planas,A.M., Bosch,A., Ferrer,I. and
Estivill,X.
A new human gene from the Down syndrome critical region encodes a
proline-rich protein highly expressed in fetal brain and heart
Hum. Mol. Genet. 4 (10), 1935-1944 (1995)
96121593
8595418
2 (bases 1 to 2346)
Fuentes,J.J., Pritchard,M.A. and Estivill,X.
Genomic organization, alternative splicing, and expression patterns
of the DSCR1 (Down syndrome candidate region 1) gene
Genomics 44 (3), 358-361 (1997)
97468152
9325060
3 (bases 1 to 2346)
Fuentes,J.J., Pritchard,M. and Estivill,X.
Direct Submission
Submitted (14-JAN-1997) Genetica Molecular, Institut de Recerca
Oncologica, Autovia de Castelldefels, km 2.7, Hospitalet, Barcelona
08907, Spain
4 (bases 1 to 2346)
Fuentes,J.J., Pritchard,M. and Estivill,X.
Direct Submission
Submitted (19-APR-2000) Genetica Molecular, Institut de Recerca
Oncologica, Autovia de Castelldefels, km 2.7, Hospitalet, Barcelona
08907, Spain
REMARK
COMMENT Sequence update by submitter
On Apr 19, 2000 this sequence version replaced gi:2612867.
FEATURES
Location/Qualifiers
source 1..2346
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.1-q22.2"
1..2346
/gene="DSCR1"
144..737
/gene="DSCR1"
/note="down syndrome candidate region 1; one of four
alternatively spliced exon 1"
/codon_start=1
/db_xref="GI:2612868"
/translation="MHFRNFYSFSLIACVANSDFSESETRAKESLPRTYDKDIT
FOYFKSPKRVIRNFSPFSAADARLQHLKTEFLGKMKLYFACTLHIGSSHLAPPND
KQFLISPPASPPVGMKQVEDATPVINYDLIYAIKSLGPKGYELHAATDTTSPVVHV
CESDQKEEEEEEMRRPKKIIQTRRPEYTHLS"
BASE COUNT 645 a 470 c 547 g 684 t
ORIGIN
Query Match 69.1%; Score 413.8; DB 9; Length 2346;
Best Local Similarity 85.1%; Pred. No. 2.4e-120;
Matches 474; Conservative 0; Mismatches 82; Indels 1; Gaps 1;
Qy 39 GAGCAGTGTTCCTTAAGCGTCTGCCCGTGAAAAAGCAGAATGATTTTAGGACTTTA 98
Db 103 GACTCGTGGTCTGTAGCGCTTTCACGTGAAGAAAGCAAGATGCAATTTAGAACTTTA 162
Qy 99 GCTACAAATTTAGCTCCCTGATTCCTTGTGTGGCAACAGATGATGCTTCAGCGAAAGTG 158
Db 163 ACTACAGTTTTCCTGCTGATTCCTTGTGTGGCAACAGATGATGCTTCAGCGAAAGTG 222
Qy 159 AGACGAGGGCCAAATTTGAATTCCTCTTCAGAACATATGACAGGACACCACTTCCAGT 218
Db 223 AAACGAGGGCCAAATTTGAGTCCCTCTTTAGGACGTATGACAGGACATCACCTTTTCAGT 282

QY 219 ATTTAAGAGCTTCAACGTCGTCCGGATAACTTACGACACCCCTTATCTGACGCCATG 278
 Db 283 ATTTAAGAGCTTCAACGAGTCAGATTAACCTTCAGCACCCCTTCCGACAGATG 342
 QY 279 CCAGGCTGGGCTGCAACAAGACCGAGTTCCTGGGGAAGAAATGAAATGTTTGGCTC 338
 Db 343 CCAGGCTCAGCTGCAATAGACTGAGTTCCTGGGGAAGAAATGAAATGTTTGGCTC 402
 QY 339 AGACTTTACATAGAAAGTTTACACCTGGCTCCGCC-AAATCCGACAAACATGTTCCCTCA 397
 Db 403 AGACTTTACATAGAAAGTTTACACCTGGCTCCGCC-AAATCCGACAAACATGTTTCCGA 462
 QY 398 TCTCCCTCCGCTCTCTCTCCGCTGGCTGGAACAGTGAAGATGCAACCCCGCTCA 457
 Db 463 TCTCCCTCCGCTCTCTCTCCGCTGGAGTGAACAGTGAAGATGCAACCCCGCTCA 522
 QY 458 TAAATTAAGATCTTTTATATGTCATCTCCAGCTGGGCGCAGAGAGAAATGAATGTC 517
 Db 523 TAACTATGATCTCTTATATGTCATCTCCAGCTGGGCGCAGAGAGAAATGAATGTC 582
 QY 518 ATGACGCGACAGACACCACTCCCAAGTGGTGGTCCACGTGTGTGAGAGTGAACAGAG 577
 Db 583 ACGACGCGACAGACACCACTCCCAAGTGGTGGTCCACGTGTGTGAGAGTGAACAGAG 642
 QY 578 ATGAGAGAGAGAGAGAA 594
 Db 643 AGAGAGAGAGAGAGAA 659

RESULT 15

AX281651 2348 bp DNA linear PAT 02-NOV-2001
 LOCUS AX281651 Sequence 60 from Patent WO0177389.
 DEFINITION AX281651
 ACCESSION AX281651
 VERSION AX281651.1 GI:16608902
 KEYWORDS

SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1
 AUTHORS Shiffman,D., Somogyi,R., Lawn,R., Seilhamer,J.J., Porter,G.J.,
 Eukaryota; Eutheria; Primates; Carnivora; Homidae; Homo.

TITLE 1
 JOURNAL Shiftman,D., Somogyi,R., Lawn,R., Seilhamer,J.J., Porter,G.J.,
 Genes expressed in foam cell differentiation
 Patent: WO 0177389-A 60 18-OCT-2001;

FEATURES
 source
 1. .2348
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /note="Incyte ID No: 042176.5"

BASE COUNT 638 a 473 c 550 g 687 t
 ORIGIN

Query Match 67.6%; Score 404.8; DB 6; Length 2348;
 Best Local Similarity 84.1%; Pred. No. 1.8e-117;
 Matches 469; Conservative 0; Mismatches 87; Indels 2; Gaps 1;

QY 39 GAGGAGCTCTTCTTAAAGCTCTGCCCCGTAAGAAAGCAAGATGATTTAGGACCTTTA 98
 Db 110 GACTGGCTGGTGTGTAGCGCTTCACTGTAAGAACCAAGATGCAATTTAGAAACTTTA 169
 QY 99 GCTTCAATTTTAACTCTCTGATGTTGTGTGGCAACGATGATGCTTCAAGCAAGATG 158
 Db 170 ACTACAGTTTAACTCTCTGATGTTGTGTGGCAACGATGATGCTTCAAGCAAGATG 229
 QY 159 AGACGAGGCGCAATTTGAATCCCTTTCAGAACATATGACAGAGACCACTTCCAGT 218
 Db 230 AAACGAGGCGCAATTTGAATCCCTTTCAGAACATATGACAGAGACCACTTCCAGT 289
 QY 219 ATTTAAGAGCTTCAACGTCGTCCGGATAACTTACGACACCCCTTATCTGACGCCATG 278

Db 290 ATTTAAGAGCTTCAACGAGTCAGATTAACCTTACGACACCCCTTCCGACAGATG 349
 QY 279 CCAGGCTGGGCTGCAACAAGACCGAGTTCCTGGGGAAGAAATGAAATGTTTGGCTC 338
 Db 350 CCAGGCTCAGCTGCAATAGACTGAGTTCCTGGGGAAGAAATGAAATGTTTGGCTC 409
 QY 339 AGACTTTACATAGAAAGTTTACACCTGG-CTCCGCCAATCCGACAAACATGTTCCCTC 396
 Db 410 AGACTTTACATAGAAAGTTTACACCTGGCTCCGCCAATCCGACAAACATGTTTCCGA 469
 QY 397 ATCTCCCTCCGCTCTCTCTCCGCTGGCTGGAACAGTGAAGATGCAACCCCGCTC 456
 Db 470 ATCTCCCTCCGCTCTCTCTCCGCTGGAGTGAACAGTGAAGATGCAACCCCGCTC 529
 QY 457 ATAAATTAAGATCTTTTATATGTCATCTCCAGCTGGGCGCAGAGAGAAATGAATGTC 516
 Db 530 ATAAATTAAGATCTTTTATATGTCATCTCCAGCTGGGCGCAGAGAGAAATGAATGTC 589
 QY 517 CATGACGCGACAGACACCACTCCCAAGTGGTGGTCCACGTGTGTGAGAGTGAACAGAG 576
 Db 590 CACGACGCGACAGACACCACTCCCAAGTGGTGGTCCACGTGTGTGAGAGTGAACAGAG 649
 QY 577 AATGAGAGAGAGAGAA 594
 Db 650 AAGAGAGAGAGAGAA 667

Search completed: December 14, 2002, 17:05:13
 Job time : 1084.63 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2002, 15:06:10 ; Search time 664.555 Seconds
(without alignments)
14597.874 Million cell updates/sec

Title: US-09-782-953-1

Perfect score: 599

Sequence: 1 gaggtcaagggaacctcca.....gaggaggaaggaagagat 599

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1:	em_estba:*
2:	em_esthum:*
3:	em_estin:*
4:	em_estnu:*
5:	em_estov:*
6:	em_estpl:*
7:	em_estro:*
8:	em_hrc:*
9:	gb_esti:*
10:	gb_est2:*
11:	gb_hrc:*
12:	gb_est3:*
13:	gb_est4:*
14:	gb_est5:*
15:	em_estfun:*
16:	em_estom:*
17:	gb_gss:*
18:	em_gss_hum:*
19:	em_gss_inv:*
20:	em_gss_pln:*
21:	em_gss_vrt:*
22:	em_gss_fun:*
23:	em_gss_man:*
24:	em_gss_mus:*
25:	em_gss_other:*
26:	em_gss_pro:*
27:	em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	599	100.0	599	9	AA200984 mul2ell.r
2	567.2	94.7	710	10	BB617325 BB617325
3	551	92.0	753	13	B1148584 602911995
4	549	91.7	980	12	BG296537 602394441
5	539.4	90.1	665	13	B1219142 602935788
6	533.8	89.1	827	13	B1102432 602889619

7	529	88.3	950	12	BF782633
8	510.8	85.3	1158	14	BQ895506
9	477	79.6	501	12	BF385073
10	461.8	77.1	711	13	BG915568
11	427.8	71.4	520	14	BQ749142
12	423	70.6	1507	11	AK010696
13	415.4	69.3	837	9	AUI24628
14	413.8	69.1	718	9	AUI31040
15	413.8	69.1	885	9	AL544755
16	413.8	69.1	939	9	AL543576
17	413.8	69.1	939	9	AL546617
18	413.8	69.1	946	14	BQ278576
19	413.8	69.1	1041	13	BM450020
20	407.8	68.1	710	12	BQ475986
21	405.8	67.7	967	14	BQ946642
22	403.8	67.4	740	13	BI463566
23	403.4	67.3	811	12	BG570239
24	402.8	67.2	662	13	BI464521
25	402.8	67.2	751	9	AL576189
26	399	66.6	906	9	AL544313
27	398.6	66.5	939	9	AL554686
28	398	66.4	1078	13	BM541636
29	386	64.4	1047	12	BG287042
30	385.4	64.3	986	12	BG828534
31	384	64.1	769	13	BI767955
32	367.8	61.4	535	14	BM832622
33	367.4	61.3	627	13	BI859506
34	364.4	60.8	525	9	AL037245
35	363.8	60.7	859	12	BF203629
36	359.4	60.0	822	12	BE895008
37	359	59.9	727	13	BI327875
38	349.4	58.3	705	9	AL550372
39	349.2	58.3	867	14	BQ427531
40	349.2	58.3	890	9	AL538796
41	348.8	58.2	828	9	AL536447
42	348.4	58.2	923	9	AL551657
43	348.2	58.1	956	9	AL556803
44	343.8	57.4	1014	9	AL559594
45	341.8	57.1	526	12	BF041330

ALIGNMENTS

RESULT 1
AA200984
LOCUS mul2ell.r1 Soares thymus 2nbmt Mus musculus cDNA clone IMAGE:619212
DEFINITION 5' similar to TR:G1125052 G1125052 DOWN SYNDROME CRITICAL REGION
PROTEIN 1. ; mRNA sequence.
AA200984
AA200984.1 GI:1795994
VERSION
KEYWORDS
SOURCE EST.
ORGANISM house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 599)
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisels,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The WashU-HMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
Washington University School
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mousees@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the

RESULT 4
 BG296537
 LOCUS 60239441P1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:4506130 5',
 DEFINITION mRNA sequence.
 ACCESSION BG296537
 VERSION BG296537.1 GI:13059271
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NIH-MGC http://mgc.nci.nih.gov/
 1 (bases 1 to 980)
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: The Cepko Laboratory
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: LLM10381 row: a column: 11
 High quality sequence start: 3
 High quality sequence stop: 747.
 Location/Qualifiers
 1..980
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone_1db="NIH_MGC_94"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: eye; Vector: pCMV-Sport6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 3.3 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH-MGC Library."
 BASE COUNT 265 a 258 c 269 g 188 t
 ORIGIN
 Query Match 91.7%; Score 549; DB 12; Length 980;
 Best Local Similarity 99.3%; Pred. No. 2.1e-155;
 Matches 593; Conservative 0; Mismatches 0; Indels 4; Gaps 4;

364 CTTGGCTCCG-CCATCCGACAAAGATTCTCTCCCTCCGCTCTCCCT 422
 Db 368 CTTGGCTCCGCCCCATCCGACAAAGATTCTCTCCCTCCGCTCTCCCT 427
 QY 423 TGGTGGAAACAGTAAAGATCCACCCCGCATTAATAGATCTTTATATGCAT 482
 Db 428 TGGTGGAAACAGTAAAGATCCACCCCGCATTAATAGATCTTTATATGCAT 487
 QY 483 CTCACAGCTGGGCGCAGAGAGAGATGATGATGATGATGATGATGATGATGATGAT 542
 Db 488 CTCACAGCTGGGCGCAGAGAGAGATGATGATGATGATGATGATGATGATGATGAT 547
 QY 543 TGGTGGTCCACCTGTTGTGAGTACCAAGATGAGAGAGAGAGAGATGATGATGAT 599
 Db 548 TGGTGGTCCACCTGTTGTGAGTACCAAGATGAGAGAGAGAGAGATGATGATGAT 604
 RESULT 5
 B1219142
 LOCUS 602935788F1 NCI_CGAP_L19 Mus musculus cDNA clone IMAGE:5093372 5',
 DEFINITION mRNA sequence.
 ACCESSION B1219142
 VERSION B1219142.1 GI:14672586
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NIH-MGC http://mgc.nci.nih.gov/
 1 (bases 1 to 665)
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: LLM11238 row: o column: 21
 High quality sequence stop: 657.
 Location/Qualifiers
 1..665
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone_1db="NCI_CGAP_L19"
 /lab_host="DH10B (TI phage-resistant)"
 /note="Organ: liver; Vector: pCMV-Sport6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.9 kb. Constructed by Life
 Technologies. Note: this is a NCI_CGAP Library."
 BASE COUNT 192 a 170 c 158 g 145 t
 ORIGIN
 Query Match 90.1%; Score 539.4; DB 13; Length 665;
 Best Local Similarity 97.6%; Pred. No. 1.4e-157;
 Matches 579; Conservative 0; Mismatches 11; Indels 3; Gaps 3;


```

Db 125 GTGCAACGATGATGCTTTTCAGCGAAAGTGAGACCA-GGCCAAATTTGAATCCCTCTTC 183
Qy 188 AGAACAATATGACAAAGGACACACCTTCCAGTATTTTAAGAGCTTCAAACGCTGTCGGGATA 247
Db 184 AGAACAATATGACAAAGGACACACCTTCCAGTATTTTAAGAGCTTCAAACGCTGTCGGGATA 243
Qy 248 AACTTCAGCAACCCCTTATCTGACGCGCATGCGAGGCTGGGCTGCAACAAGCCGAGTTC 307
Db 244 AACTTCAGCAACCCCTTATCTGACGCGCATGCGAGGCTGGGCTGCAACAAGCCGAGTTC 303
Qy 308 CTGGGGAAGGAAATGAAGTTGTATTTTGTCTCAGACTTTTACACATAGGAAGTTCCACCTG 367
Db 304 CTGGGGAAGGAAATGAAGTTGTATTTTGTCTCAGACTTTTACACATAGGAAGTTCCACCTG 363
Qy 368 GCTCCG-CCAAATCCGCAACAAAGTTCCTCATCTCCCTCCGCTCTCCCTCCGTTGGC 426
Db 364 GCTCCGCCCAATCCGCAACAAAGTTCCTCATCTCCCTCCGCTCTCCCTCCGTTGGC 423
Qy 427 TGGAAACAAGTAGACCAATTTATACCCCGTCAATAAATTACGATCTTTTATATGCCATCTCC 486
Db 424 TGGAAACAAGTAGACCAATTTATACCCCGTCAATAAATTACGATCTTTTATATGCCATCTCC 483
Qy 487 AAGCTGGGGCCAGGAGAGAGTATGAATGCTGATGCGAGCAGACACCACTCCCAAGTGTG 546
Db 484 AAGCTGGGGCCAGGAGAGAGTATGAATGCTGATGCGAGCAGACACCACTCCCAAGTGTG 543
Qy 547 GTGGTCCACGTGTGTGAGAGTGACCAAGAGAAATGAGGAGGAAGGAAGAGAT 599
Db 544 GTGGTCCACGTGTGTGAGAGTGACCAAGAGAAATGAGGAGGAAGGAAGAGAT 596

RESULT 6
BI102432
LOCUS 602889619F1 NCI CGAP_Kid14 Mus musculus cDNA clone IMAGE:5044612
DEFINITION 5', mRNA sequence.
ACCESSION BI102432
VERSION BI102432.1 GI:14553325
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Contact: Robert Strausberg, Ph.D.
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
plate: LLAM1122 row: f column: 05
High quality sequence stop: 666.
Location/Qualifiers
1. .827
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5044612"
/lab_host="NCI CGAP_Kid14"
/note="Organ: kidney; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.75 kb. Constructed by Life
Technologies. Note: this is a NCI CGAP Library. |"
206 a 248 c 204 g 169 t

```

```

Query Match 89.1%; Score 533.8; DB 13; Length 827;
Best Local Similarity 98.8%; Pred. No. 7.8e-151;
Matches 590; Conservative 0; Mismatches 2; Indels 5; Gaps 5;

Qy 6 GCAAGGAACCTCCAGCTTGGGCTTTGACTGAGAGCGAGTCTGTTTAAAGCGTCTGCC 65
Db 1 GCAAGGAACCTCCAGCTTGGGCTTTGACTGAGAGCGAGTCTGTTTAAAGCGTCTGCC 60
Qy 66 CCGTGAAGAAGCAGATG-ATTTTAGGAGCTTTAGTACAATTTTAGCTCCCTGATTGT 124
Db 61 CCGTGAAGAAGCAGATGCAATTTTAGGAGCTTTAGTACAATTTTAGCTCCCTGATTGT 120
Qy 125 TGTGTGGCAAAACGATGATGCTTTAGCGGAAAGTGAGACCAGGCGCCAAATTTTCAATCCCTC 184
Db 121 TGTGTGGCAAAACGATGATGCTTTAGCGGAAAGTGAGACCA-GGCCAAATTTTGAATCCCTC 179
Qy 185 TTCAGAACATATGACAAAGGACACCACTTCCAGTATTTTAAGAGCTTCAAACGTTCCCG 244
Db 180 TTCAGAACATATGACAAAGGACACCACTTCCAGTATTTTAAGAGCTTCAAACGTTCCCG 239
Qy 245 ATAACTTCAGCAACCCCTTATCTGACGCGATGCGAGGCTGGGCTGCAACAAGACCGAG 304
Db 240 ATAACTTCAGCAACCCCTTATCTGACGCGATGCGAGGCTGGGCTGCAACAAGACCGAG 299
Qy 305 TTCTCTGGGGAAGAAATGAAGTTGTATTTTGTCTCAGACTTTTACACATAGGAAGTTCAAC 364
Db 300 TTCTCTGGGGAAGAAATGAAGTTGTATTTTGTCTCAGACTTTTACACATAGGAAGTTCAAC 359
Qy 365 CTGGCTCCG-CCAAATCCCGCAACAAAGTTCCTCATCTCCCTCCGCTCTCTCCCTCCGTT 423
Db 360 CTGGCTCCGCGCAATCCCGCAACAAAGTTCCTCATCTCCCTCCGCTCTCTCCCTCCGTT 419
Qy 424 GGCTGAAACAAAGTAGAAGATGCCACCCCTCATAAATTTAGATCTTTTATATGCCATC 483
Db 420 GGCTGAAACAAAGTAGAAGATGCCACCCCTCATAAATTTAGATCTTTTATATGCCATC 479
Qy 484 TCCAAAGCTGGGCGCCAGGAGAGAGTATGAATGCTGATGCGAGGACAGACACCACTCCCACT 543
Db 480 TCCAAAGCTGGGCGCCAGGAGAGAGTATGAATGCTGATGCGAGGACAGACACCA-TCCTCA 538
Qy 544 GTGGTGGTCCAGTGTGTGAGAGTGACCAAGAGAAATGAGG-AGGAAGAGGAAGAGAT 599
Db 539 GTGGTGGTCCAGTGTGTGAGAGTGACCAAGAGAAATGAGGAGGAGGAGGAAGAGAT 595

RESULT 7
BI102432
LOCUS 602107462F1 NCI CGAP_Kid14 Mus musculus cDNA clone IMAGE:4236038
DEFINITION 5', mRNA sequence.
ACCESSION BI102432
VERSION BI102432.1 GI:12087669
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Contact: Robert Strausberg, Ph.D.
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: LLAM9843 row: k column: 15
High quality sequence stop: 701.

```


BF385073 501 bp mRNA linear EST 27-NOV-2000
LOCUS 602045671P1 NCI_CGAP_Li9 Mus musculus cDNA clone IMAGE:4195384 5',
DEFINITION mRNA sequence.
ACCESSION BF385073
VERSION BF385073.1 GI:11366365
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 501)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AMP530 row: m column: 17
High quality sequence stop: 501.
Location/Qualifiers
1. 501
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4195384"
/clone_lib="NCI_CGAP_Li9"
/lab_host="DHI08 (T1 phage-resistant)"
/notes="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.9 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 132 a 134 c 113 g 122 t
ORIGIN
Query Match 79.6%; Score 477; DB 12; Length 501;
Best Local Similarity 99.6%; Pred. No. 1e-133;
Matches 499; Conservative 0; Mismatches 0; Indels 2; Gaps 2;
QY 16 CTCAGCTGGCTGACTGAGAGAGCGAGTCTGTTAAAGCGTGTGCCCGTGAAGAAA 75
DB 1 CTCAGCTGGCTGACTGAGAGAGCGAGTCTGTTAAAGCGTGTGCCCGTGAAGAAA 60
QY 76 GCAGATG-ATTTAGGACTTTAGCTACAAATTTAGCTCCCTGATGCTGTGTGGCAA 134
DB 61 GCAGATGCAATTTAGGACTTTAGCTACAAATTTAGCTCCCTGATGCTGTGTGGCAA 120
QY 135 ACATGATGCTTCAGCGAAGTGAGACCGAGGCGCAAAATTTGAATCCCTCTTCAGAACAT 194
DB 121 ACATGATGCTTCAGCGAAGTGAGACCGAGGCGCAAAATTTGAATCCCTCTTCAGAACAT 180
QY 195 ATCAAGAGACACACCTTCAGTATTTAAAGAGCTTCAAACTGTGCGGATAAACTTCA 254
DB 181 ATCAAGAGACACACCTTCAGTATTTAAAGAGCTTCAAACTGTGCGGATAAACTTCA 240
QY 255 GCAACCCCTTATCTGAGCGGATGCGAGGCTGCGGCTGCACAGACCGAGTTCCTGGGGA 314
DB 241 GCAACCCCTTATCTGAGCGGATGCGAGGCTGCGGCTGCACAGACCGAGTTCCTGGGGA 300
QY 315 AGAAATGAAGTTGATTTTGTCTCAGACTTTACATAGGAAAGTTTACACCTGGCTCCG- 373
DB 301 AGAAATGAAGTTGATTTTGTCTCAGACTTTACATAGGAAAGTTTACACCTGGCTCCG- 360
QY 374 CCAATCCGACAAACAGTTTCTCATCTCCCTCCGCGCTCTCTCCGTTGGCTGGAAC 433
DB 361 CCAATCCGACAAACAGTTTCTCATCTCCCTCCGCGCTCTCTCCGTTGGCTGGAAC 420

QY 434 AACTAGAGATGCCACCCCGTCATAAATACGATCTTTTATATGCGCATCTCCAAGCTGG 493
DB 421 AACTAGAGATGCCACCCCGTCATAAATACGATCTTTTATATGCGCATCTCCAAGCTGG 480
QY 494 GCCAGGAGAGAGTATGAAC 514
DB 481 GCCAGGAGAGAGTATGAAC 501
RESULT 10
BG915568
LOCUS 602815815F1 NCI_CGAP_Mam4 Mus musculus cDNA clone IMAGE:4945346 5',
DEFINITION mRNA sequence.
ACCESSION BG915568
VERSION BG915568.1 GI:14296044
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 711)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth
Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM10891 row: n column: 03
High quality sequence stop: 711.
Location/Qualifiers
1. 711
/organism="Mus musculus"
/strain="NMRI"
/db_xref="taxon:10090"
/clone="IMAGE:4945346"
/clone_lib="NCI_CGAP_Mam4"
/tissue_type="tumor_gross tissue"
/dev_stage="5 months"
/lab_host="DHI08"
/notes="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Priscilla Furth,
NIH Reference for transgenic model: Li et al., Cell Growth
and Differentiation 7, 3-11 (1996)."
BASE COUNT 196 a 184 c 183 g 148 t
ORIGIN
Query Match 77.1%; Score 461.8; DB 13; Length 711;
Best Local Similarity 99.0%; Pred. No. 5.3e-129;
Matches 496; Conservative 0; Mismatches 2; Indels 3; Gaps 3;
QY 101 TACAATTTAGCTCCCTGATGCTGTGTGGCAACCATGATGCTTTCAGCGAAAGTGAG 160
DB 1 TACAATTTAGCTCCCTGATGCTGTGTGGCAACCATGATGCTTTCAGCGAAAGTGAG 60
QY 161 ACCAGGCGCAAAATTTGAATCCCTCTTCAGAACATATGACAGGACACCACTTCCAGTAT 220
DB 61 ACCA-GGCCAAATTTGAATCCCTCTTCAGAACATATGACAGGACACCACTTCCAGTAT 119
QY 221 TTTAAGAGCTTCAAACTGTGCGGATAAACTTCAGCAACCCCTTATCTGAGCGGATGCC 280
DB 120 TTTAAGAGCTTCAAACTGTGCGGATAAACTTCAGCAACCCCTTATCTGAGCGGATGCC 179
QY 281 AGGCTGGCGTGCACAGACCGAGTTCTCTGGGGAAGGAAATGAAGTTGTATTGTTGCTCAG 340

```

Db 180 AGCTGCGGCTCACAAGCCAGTTCCTGGGAGAAATGAATTGTTCTCTCAG 239
Qy 341 ACTTACACATGAGAGATTGACACCTGCTCCG-CCATCCCGACAAACAGTTCTCATC 399
Db 240 ACTTACACATGAGAGATTGACACCTGCTCCGCGCCCAATCCGCAACACAGTTCTCATC 299
Qy 400 TCCCTCCCGGCTCT-CCCTCCGTTGGCTGGAACAGTAGAAGTGCACCCCGTCAT 458
Db 300 TCCCTCCCGGCTCTCTCCCTCCGCTGGGCTGAAACAGTAGAAGTGCACCCCGTCAT 359
Qy 459 AATTACGATCTTTATATATGATCTCCAGTGGGCGCAGAGAGAACTATGAAGTCA 518
Db 360 AATTACGATCTTTATATATGATCTCCAGTGGGCGCAGAGAGAACTATGAAGTCA 419
Qy 519 TCGAGCAGACAGACCACTCCAGTGTGTGTCCAGTGTGTGAGAGACCAAGAGA 578
Db 420 TCGAGCAGACAGACCACTCCAGTGTGTGTCCAGTGTGTGAGAGACCAAGAGA 479
Qy 579 TGAGGAGAGAGAGAGAGAT 599
Db 480 TGAGGAGAGAGAGAGAGAT 500

RESULT 11
LOCUS B0749142 520 bp mRNA linear EST 17-JUL-2002
DEFINITION UI-M-FD0-bvd-c-04-0-UI.r1 NIH_BMAP_FD0 Mus musculus cDNA clone
IMAGE:5716539 5', mRNA sequence.
ACCESSION B0749142
VERSION B0749142.1 GI:21895929
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 520)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs.fda.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5.
Location/Qualifiers
1. 520
/organism="Mus musculus"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:5716539"
/clone_1lb="NIH BMAP_FD0"
/clone_type="Whole Brain"
/dev_stage="embryo 12.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/note="Organ: Brain; Vector: pYX-Asc; Site: 1: Ecor I;
Site 2: Not I; The library was constructed according to
Bonaldi, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured mRNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was size selected according to mRNA size fraction,
ligated with Ecor I adaptor, digested with Not I, and then
cloned directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is TGAGGAGAGCC. This library was created for the

```

```

BASE COUNT 137 a 147 c 131 g 105 t
ORIGIN
Query Match 71.4%; Score 427.8; DB 14; Length 520;
Best Local Similarity 93.3%; Pred. No. 9.3e-119;
Matches 458; Conservative 0; Mismatches 32; Indels 1; Gaps 1;

University of Iowa Mouse Brain Molecular Anatomy Project
(BMAP): 'Gene Discovery in the Developing Mouse Nervous
System', supported by National Institutes of Mental Health
(NIMH), Hemin Chin, Ph.D., program coordinator."

Qy 110 AGCTCCGATGTTGTTGAGCAAAAGATGATGTTCCAGCAAAAGTAGACGAGGCC 169
Db 12 AGCGCCACCATGCGCTCCACCTGGACCCCGGCTGTTCTGAGCGGCTGTCCGGGCC 71
Qy 170 AATTGGAATCCCTCTCCAGAACATATGACAGACACCACTTCAGTATTTTAAGAGC 229
Db 72 AATTGGAATCCCTCTCCAGAACATATGACAGACACCACTTCAGTATTTTAAGAGC 131
Qy 230 TTCAACGTTGCGGATTAACCTTACGCAACCCCTTATCTGACGCCATGCGCAGGTCGG 289
Db 132 TTCAACGTTGCGGATTAACCTTACGCAACCCCTTATCTGACGCCATGCGCAGGTCGG 191
Qy 290 CTGCAAGAGACGAGTTCCTGGGAGAAATGAAGTTGTTTGTCTCAGACTTTACAC 349
Db 192 CTGCAAGAGACGAGTTCCTGGGAGAAATGAAGTTGTTTGTCTCAGACTTTACAC 251
Qy 350 ATAGGAATTCACACCTGCTCCG-CCATCCGACAAACAGTTCATCTCCCTCCG 408
Db 252 ATAGGAATTCACACCTGCTCCGCGCCCATCCGACAAACAGTTCATCTCCCTCCG 311
Qy 409 GCCTTCTCTCCGTTGCTGGAACCAAGTAGAATGCCACCCCGCATTAATTAAGAT 468
Db 312 GCCTTCTCTCCGTTGCTGGAACCAAGTAGAATGCCACCCCGCATTAATTAAGAT 371
Qy 469 CTTTATATGCAATCTCCAGCTGGGCGCAGAGAGAAATGAATGATGATGACGACGCA 528
Db 372 CTTTATATGCAATCTCCAGCTGGGCGCAGAGAGAAATGAATGATGACGACGCA 431
Qy 529 GACACCACTCCCAAGTGTGTGCTCCAGTGTGTGAGTACCAAGAAATGAGAGAGA 588
Db 432 GACACCACTCCCAAGTGTGTGCTCCAGTGTGTGAGTACCAAGAAATGAGAGAGA 491
Qy 589 GAGGAAGAGAT 599
Db 492 GAGGAAGAGAT 502

RESULT 12
LOCUS AK010696 1507 bp mRNA linear HTC 19-JAN-2002
DEFINITION AK010696 Mus musculus ES cells cDNA, RIKEN full-length enriched library,
clone:2410048A02:Down syndrome critical region homolog 1 (human),
full insert sequence.
ACCESSION AK010696
VERSION AK010696.1 GI:12846329
KEYWORDS HTC, CAP trapper.
SOURCE Mus musculus (strain: C57BL/6J) ES cells cDNA to mRNA,
clone_1lb:RIKEN full-length enriched mouse cDNA library
clone:2410048A02.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 Carninci, P. and Hayashizaki, Y.
AUTHORS High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE
2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
AUTHORS Normalization and subtraction of cap-trapper-selected cDNAs to
TITLE

```


LOCUS AU124628 837 bp mRNA linear EST 01-AUG-2002
 DEFINITION AU124628 NT2RM4 Homo sapiens CDNA clone NT2RM4000307 5', mRNA
 ACCESSION AU124628
 VERSION AU124628.1 GI:10949344
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Ota,T., Wakamatsu,A., Ozawa,M., Ishii,S., Saito,K., Yamamoto,J., Nakamura,Y., Nishikawa,T., Nagai,T., Suzuki,Y., Sugano,S. and Iisogai,T.
 HRI human CDNA project (Ota,T., Wakamatsu,A., Ozawa,M., Ishii,S., Saito,K., Yamamoto,J., Nakamura,Y., Nishikawa,T., Nagai,T., Suzuki,Y., Sugano,S., Iisogai,T., Sugano,S., Iisogai,T.)
 Unpublished (2000)
 JOURNAL Contact: Takao Iisogai
 Genomics Laboratory
 Helix Research Institute
 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
 Tel: 81-438-52-3975
 Fax: 81-438-52-3986
 Email: genomice@hri.co.jp
 HRI human CDNA project; 5'- & 3'-end one pass sequencing; Helix Research Institute; CDNA library construction; Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.
 Location/Qualifiers
 1..837
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_id="NT2RM4000307"
 /clone_1lb="NT2RM4"
 /cell_type="teratocarcinoma"
 /note="Vector: pME18FLJ; mRNA from uninduced NT2 neuronal precursor cells"
 BASE COUNT 241 a 201 c 207 g 186 t 2 others
 ORIGIN
 Query Match 69.3%; Score 415.4; DB 9; Length 837;
 Best Local Similarity 85.3%; Pred. No. 7.4e-115;
 Matches 475; Conservative 0; Mismatches 81; Indels 1; Gaps 1;
 Oy 39 GAGCGAGTCTGTTAGGCTCTGCGCGGAAAGAGAAATGATTTAGGACTTTA 98
 Db 77 GACTGCGTGGTCTGTAGCGCTTCACTGTAGAAAGCAAGATGCAATTTAGAACTTTA 136
 Oy 99 GCTACATTTTAACTCCCTGATTTGCTGTGTGCAACGATGATGCTTACGCGAAAGTG 158
 Db 137 ACTACAGTTTAACTCCCTGATTTGCTGTGTGCAACGATGATGCTTACGCGAAAGTG 196
 Oy 159 AGACGAGGCGCAATTTGAATCCCTCTTCAGAACATATGACAAAGAACACACCTTCCAGT 218
 Db 197 AAACGAGGCGCAATTTGAATCCCTCTTCAGAACATATGACAAAGAACATATGATGAT 256
 Oy 219 ATTTTAAGAGCTTCAAGCGTCCGAGTAAACTTCAGAACCCCTTATCTGACCGCATG 278
 Db 257 ATTTTAAGAGCTTCAAGCGATCAGAAATTAATTCAGAACCCCTTCTCCGACGCAATG 316
 Oy 279 CCAGGCTGCGGCTGCACAAAGACCGAGTTCTCTGGGAAAGAAATGAAGTTGATTTTCTC 338
 Db 317 CCAAGCTCCAGCTGCATAGACTGAGTTCTGGGAAAGAAATGAAGTTATATTTGCTC 376
 Oy 339 AGACTTACATAGAAAGTTTCAACCTGCTCCGCC-AAATCCGACAAACAGTTCTCTCA 397
 Db 377 AGACTTACATAGAAAGTTTCAACCTGCTCCGCCAAATCCGACAAACAGTTCTCTCA 436
 Oy 398 TCTCCCTCCGCGCTCTCTCCGTTGGCTGAAAAAGATGAGATGCCACCCCGCTCA 457
 Db 437 TCTCCCTCCGCGCTCTCTCCGCGAGTGGAGTGAACAAAGTGAAGATGCCACCCAGTCA 496

RESULT 14
 LOCUS AU131040 718 bp mRNA linear EST 01-AUG-2002
 DEFINITION AU131040 NT2RFP3 Homo sapiens CDNA clone NT2RFP3001895 5', mRNA
 ACCESSION AU131040
 VERSION AU131040.1 GI:10991394
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and Iisogai,T.
 HRI human CDNA project
 Unpublished (2000)
 JOURNAL Contact: Takao Iisogai
 Genomics Laboratory
 Helix Research Institute
 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
 Tel: 81-438-52-3975
 Fax: 81-438-52-3986
 Email: genomice@hri.co.jp
 HRI human CDNA project; 5'- & 3'-end one pass sequencing; Helix Research Institute; CDNA library construction; Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.
 Location/Qualifiers
 1..718
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_id="NT2RFP3001895"
 /clone_1lb="NT2RFP3"
 /cell_type="teratocarcinoma"
 /note="Vector: pME18FLJ; mRNA from NT2 neuronal precursor cells after 2-weeks retinoic acid (RA) induction"
 BASE COUNT 217 a 174 c 169 g 155 t 3 others
 ORIGIN
 Query Match 69.1%; Score 413.8; DB 9; Length 718;
 Best Local Similarity 85.1%; Pred. No. 2e-114;
 Matches 474; Conservative 0; Mismatches 82; Indels 1; Gaps 1;
 Oy 39 GAGCGAGTCTGTTAGGCTCTGCGCGGAAAGAGAAATGATTTAGGACTTTA 98
 Db 61 GACTGCGTGGTCTGTAGCGCTTCACTGTAGAAAGCAAGATGCAATTTAGAACTTTA 120
 Oy 99 GCTACATTTTAACTCCCTGATTTGCTGTGTGCAACGATGATGCTTACGCGAAAGTG 158
 Db 121 ACTACAGTTTAACTCCCTGATTTGCTGTGTGCAACGATGATGCTTACGCGAAAGTG 180
 Oy 159 AGACGAGGCGCAATTTGAATCCCTCTTCAGAACATATGACAAAGAACACACCTTCCAGT 218
 Db 181 AAACGAGGCGCAATTTGAATCCCTCTTCAGAACATATGACAAAGAACATATGATGAT 240
 Oy 219 ATTTTAAGAGCTTCAAGCGTCCGAGTAAACTTCAGAACCCCTTATCTGACCGCATG 278

```

Db 241 *ATTTTAAGAGCTTCAACAGAGTGCAGATTAACATTAACCTTCAGCAACCCCTTCTCCGACGAGATG 300
QY 279 CAGAGTGGCGGTGCACAAGACCGAGTTCCTGGGGAAGGAATGAAGTTGATTTTGGTTC 338
Db 301 CAGAGTCCAGCTGCATAAGAGCTGAGTTCTGGGAAGGAATGAAGTTATATTTTGGTTC 360
QY 339 AGACTTTACACATAGGAAGTTACACCTGGCTCCGCGC-AATCCGACAAACAGTTTCCTCA 397
Db 361 AGACCTTTACACATAGGAAGCTCACACCTGGCTCCGCGCAATCCAGCAAGCAGTTTCTGA 420
QY 398 TCTCCCTCCGGCTCTCTCCGCTGGCTGGGAACAAGTAGAAGTGCACCCCGCTCA 457
Db 421 TCTCCCTCCGGCTCTCCGCTGGCTGGGAACAAGTAGAAGTGCACCCCGCTCA 480
QY 458 TAAATTAAGAGCTTCAACAGAGTGCAGATTAACATTAACCTTCAGCAACCCCTTCTCCGACGAGATG 517
Db 481 TAAACTATGATCTTATATGCATCTCCAGCTGGGCGGCAAGGAATGAAGTTGATTTTGGTTC 540
QY 518 ATGCAGCGACAGACCACTCCAGCTGGTGGTCCAGCTGGGCGGCAAGGAATGAAGTTGATTTTGGTTC 577
Db 541 ACGCAGCGACTGACACCACTCCAGCTGGTGGTCCAGCTGGGCGGCAAGGAATGAAGTTGATTTTGGTTC 600
QY 578 ATGAGAGGAAGAGGAA 594
Db 601 AGGAGGAAGAGAGGAA 617

RESULT 15
LOCUS AL544755 885 bp mRNA linear EST 16-FEB-2001
DEFINITION AL544755 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0D1022YJ23 5
prime, mRNA sequence.
ACCESSION AL544755
VERSION AL544755.1 GI:12877235
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
source
1..885
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1022YJ23"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/notes="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT)-primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 247 a 211 c 228 g 198 t 1 others
ORIGIN

Query Match 69.1%; Score 413.8; DB 9; Length 885;
Best Local Similarity 85.1%; Pred. No. 2.3e-114;
Matches 474; Conservative 0; Mismatches 82; Indels 1; Gaps 1;

QY 39 GAGCGAGTCGTTGTTAAGCGTCTGCCCGTGGAAAGCAGAATGATTTTAGGACTTTA 98

```

```

Db 9 GACTGCGTGGGTCTGTAGCGCTTTCACTGTAAAGAAAGCAAGATGCAATTTTAGAACTTTA 68
QY 99 GCTACAAATTTTAGCTCCCTGATTGCTTGTGTGGCAACGATGATGCTTCAGCGAAAGTG 158
Db 69 ACTACAGTTTTAGCTCCCTGATTGCTTGTGTGGCAACGATGATGCTTCAGCGAAAGTG 128
QY 159 AGACCAGGGGCAAAATTTGAATCCCTCTTTCAGAACATATGACAAAGGACATCCACCTTCAGT 218
Db 129 AAACCAAGGGGCAAAATTTGAGTCCCTCTTTAGGACGTATGACAAAGGACATCCACCTTCAGT 198
QY 219 ATTTTAAGAGCTTCAACAGCTGTCGGGATAACTTTCAGCAACCCCTTATCTGCAAGCCGATG 278
Db 189 ATTTTAAGAGCTTCAACAGCTGTCAGATTAACCTTTCAGCAACCCCTTCTCCGACGAGATG 248
QY 279 CAGGCTGGGGTGCACAAGACCGAGTTCCTGGGGAAGGAATGAAGTTGATTTTGGTTC 338
Db 249 CCAGGCTCCAGCTGCATTAAGACTGAGTTTCTGGGAAGGAATGAAGTTATATTTTGGTTC 308
QY 339 AGACTTTACACATAGGAAGTTACACCTGGCTCCGCGC-AATCCGACAAACAGTTTCCTCA 397
Db 309 AGACCTTTACACATAGGAAGCTCACACCTGGCTCCGCGCAATCCAGCAAGCAGTTTCTCA 368
QY 398 TCTCCCTCCGGCTCTCTCCCTCCGCTGGCTGGAAACAAGTAGAAGTGCACCCCGCTCA 457
Db 369 TCTCCCTCCGGCTCTCTCCGCTGGATGGAAACAAGTAGAAGTGCACCCCGCTCA 428
QY 458 TAAATTAAGAGCTTCAACAGCTGTCGGGCGGCAAGGAATGAAGTTGATTTTGGTTC 517
Db 429 TAACTATGATCTTATATGCATCTCCAGCTGGGCGGCAAGGAATGAAGTTGATTTTGGTTC 488
QY 518 ATGCAGCGACAGACCACTCCAGCTGGTGGTCCAGCTGGGCGGCAAGGAATGAAGTTGATTTTGGTTC 577
Db 489 ACGCAGCGACTGACACCACTCCAGCTGGTGGTCCAGCTGGGCGGCAAGGAATGAAGTTGATTTTGGTTC 548
QY 578 ATGAGAGGAAGAGGAA 594
Db 549 AGGAGGAAGAGAGGAA 565

```

Search completed: December 14, 2002, 18:10:49
Job time : 670.555 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2002, 14:10:15 ; Search time 380.679 Seconds
(without alignments)
13789.580 Million cell updates/sec

Title: US-09-782-953-11

Perfect score: 2331

Sequence: 1 tttttttttccccaggaggt.....gggcacagtagtctccttac 2331

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2331	100.0	2331	24	Human MCIP associa
2	2318	4	2348	24	Human DNA sequence
3	2109	90.5	2212	24	Human MCIP associa
4	2101	90.1	2173	24	Gene #3341 used to
5	2094.2	89.8	2358	21	Lung cancer associ
6	2059.2	88.3	2174	20	Human DSCR1 coding
7	2059.2	88.3	2174	20	Human breast cell
8	486	20.8	486	22	Human foetal liver
9	486	20.8	486	22	Probe #809 for gen

C 10	486	20.8	486	22	AAK00816	Human brain expres
C 11	486	20.8	486	22	AAK26270	Human bone marrow
C 12	486	20.8	486	22	AAI10903	Probe #836 for gen
C 13	486	20.8	486	22	AAI32163	Probe #849 used to
C 14	486	20.8	486	22	AAI00825	Probe #816 used to
C 15	486	20.8	486	22	ABS00858	Human genome-deriv
C 16	470.6	20.2	597	24	AAQ30153	Mouse MCIP associa
C 17	442	19.0	442	21	AAQ30153	Human MCIP associa
C 18	413.8	17.8	599	24	AAQ30153	Human secreted pro
C 19	403.8	17.3	597	24	AAQ30153	Human MCIP associa
C 20	382.4	16.4	404	22	AAI23689	Human breast tumor
C 21	365.2	15.7	385	22	AAH55546	Human breast cancer
C 22	327.6	14.1	640	24	ABQ59719	Human breast tumor
C 23	246.8	10.6	251	20	AAH55546	DNA encoding human
C 24	246.8	10.6	251	21	AAH55546	Known human breast
C 25	246.8	10.6	251	24	ABK29002	Human breast tumor
C 26	194.4	8.3	3184	24	AAQ30153	Human MCIP associa
C 27	194.4	8.3	3184	24	ABL61768	Colon adenocarcino
C 28	194.4	8.3	3184	24	ABL61768	Lung cancer relate
C 29	193.2	8.3	594	24	AAQ30153	Human MCIP associa
C 30	189	8.1	720	24	AAQ30153	Human MCIP associa
C 31	189	8.1	828	24	AAQ30153	Human MCIP associa
C 32	187	8.0	934	22	AAF25338	Nucleotide sequenc
C 33	187	8.0	3159	24	ABA91457	Human Down syndrom
C 34	176.2	7.6	415	22	ABL14827	Human breast cancer
C 35	175	7.5	412	22	ABA56172	Human foetal liver
C 36	175	7.5	412	22	AAK04364	Human brain expres
C 37	175	7.5	412	22	AAI14449	Probe #4382 for ge
C 38	175	7.5	412	22	AAI35821	Probe #4507 used t
C 39	175	7.5	412	22	AAI04272	Probe #4263 used t
C 40	175	7.5	412	24	ABS04419	Human genome-deriv
C 41	175	7.5	446	22	ABA34406	Human breast cell
C 42	175	7.5	446	22	ABA33852	Human foetal liver
C 43	175	7.5	446	22	ABA23598	Probe #2064 for ge
C 44	175	7.5	446	22	AAK02113	Human brain expres
C 45	175	7.5	446	22	AAK27563	Human bone marrow

ALIGNMENTS

RESULT 1

AAQ30155
ID AAD30155 standard; DNA; 2331 BP.

AC AAD30155;

XX 17-MAY-2002 (first entry)

DT Human MCIP associated DNA #1.

DE Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;

KW heart failure; cardiomyopathy; heart disease; human; gene; ds.

XX Homo sapiens.

OS Homo sapiens.

XX Key

FT CDS

FT Location/Qualifiers

FT 144...737

FT /*tag= a

FT /product= "Human MCIP associated protein #1"

XX WO200204491-A2.

PN 17-JAN-2002.

PD 06-JUL-2001; 2001WO-US21662.

XX 07-JUL-2000; 2000US-216601P.

PR 13-FEB-2001; 2001US-0782953.

XX (TEXA) UNIV TEXAS SYSTEM.

PA (WILL/) WILLIAMS S R.

PA (ROTH/) ROTHERMEL B.

XX Williams SR, Rothenmel B;
 XX WPI: 2002-179698/23.
 DR P-PSDB: AAE18914.
 XX
 XX Screening for modulators of muscle calcineurin interacting protein
 PT (MCIP) binding, expression or phosphorylation, useful for treating
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,
 PT calcineurin and a test compound -
 XX
 XX Claim 72: Page 155-157; 174pp; English.
 XX
 XX The invention relates to muscle calcineurin interacting proteins (MCIPs)
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical
 CC complex with the catalytic subunit of calcineurin and increased levels
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate
 CC transcription of certain target genes. The invention also relates to
 CC methods for identifying modulators of MCIP binding, expression or
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin
 CC may be used for treating cardiac hypertrophy and heart failure.
 CC Antibodies to MCIP can be used in characterizing the MCIP content of
 CC healthy and diseased tissues and subsequently for determining the
 CC presence or absence of cardiomyopathy or as predictor of heart disease.
 CC The present sequence is human MCIP associated DNA.
 CC Note: This sequence has been described as a promoter in claim 72 of
 CC the specification, however the sequence seems to be a polynucleotide
 CC encoding a MCIP associated protein.
 XX
 XX Sequence 2331 BP; 630 A; 470 C; 547 G; 684 T; 0 other;
 SO

Query Match 100.0%; Score 2331; DB 24; Length 2331;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTTTTCCTCCAGAGAGTGGGGCTGGCCCTTACTGCTTTATTAAGACCCAGCTCAAA 60
 Db 1 TTTTTCCTCCAGAGAGTGGGGCTGGCCCTTACTGCTTTATTAAGACCCAGCTCAAA 60
 OY 61 AGGACCTACAGCCTCTTGGAAGGAATCTCATAGGGGCTGACGGTGGCTGTAG 120
 Db 61 AGGACCTACAGCCTCTTGGAAGGAATCTCATAGGGGCTGACGGTGGCTGTAG 120
 OY 121 CGCTTCACCTGTAAGAAAGCAATGATCTTTAGAACTTTAACTACAGTTTACCTCC 180
 Db 121 CGCTTCACCTGTAAGAAAGCAATGATCTTTAGAACTTTAACTACAGTTTACCTCC 180
 OY 181 TGAATGCTGTGGCAAGATGATATCTTCAGCGAAAGTGAACAGGGCCNAATTTG 240
 Db 181 TGAATGCTGTGGCAAGATGATATCTTCAGCGAAAGTGAACAGGGCCNAATTTG 240
 OY 241 AGTCCCTCTTGAAGACTATGACAAAGCATCACTTTAGATTTTAAAGCTTCAAC 300
 Db 241 AGTCCCTCTTGAAGACTATGACAAAGCATCACTTTAGATTTTAAAGCTTCAAC 300
 OY 301 GAGTCAGATTAATCTTCAAGCAACCTTCTCCGACAGATGCCAGGCTCAAGCTGCATA 360
 Db 301 GAGTCAGATTAATCTTCAAGCAACCTTCTCCGACAGATGCCAGGCTCAAGCTGCATA 360
 OY 361 AGACTGAGTTTCTGGGAAAGGAAATGAATATATTTTGTCTCAACCTTACACATAGAAA 420
 Db 361 AGACTGAGTTTCTGGGAAAGGAAATGAATATATTTTGTCTCAACCTTACACATAGAAA 420
 OY 421 GCTACACCTGGCTCCGCAAAATCCAGCAAGAGTTTCTGATCTCCCTCCGCTCTC 480
 Db 421 GCTACACCTGGCTCCGCAAAATCCAGCAAGAGTTTCTGATCTCCCTCCGCTCTC 480
 OY 481 CGCCAGTGGATGGAACCAAGTGAAGATGACCCCAAGTCATAAATCTATAT 540
 Db 481 CGCCAGTGGATGGAACCAAGTGAAGATGACCCCAAGTCATAAATCTATAT 540
 OY 541 ATGCCATCTCAAGCTGGGGCCAGGGGAAAGTATGAAATTCAGACGAGCTGACACA 600
 Db 541 ATGCCATCTCAAGCTGGGGCCAGGGGAAAGTATGAAATTCAGACGAGCTGACACA 600

Db 541 ATGCCATCTCAAGCTGGGGCCAGGGGAAAGTATGAAATTCAGACGAGCTGACACA 600
 OY 601 CTCCAGCGTGGTGTGCATATGATGAGATGATCAAGAGAGAGAGAGAGAAA 660
 Db 601 CTCCAGCGTGGTGTGCATATGATGAGATGATCAAGAGAGAGAGAGAGAAA 660
 OY 661 TGGAAAGATGAGAGAGACTTAAGCCAAAATTTATCCAGACCGAGGCGCGAGTACGCC 720
 Db 661 TGGAAAGATGAGAGAGACTTAAGCCAAAATTTATCCAGACCGAGGCGCGAGTACGCC 720
 OY 721 CGATTCACCTGAGCTGACCTGGACGCGAGAGAGCGACTTCCAAATCTACACGGA 780
 Db 721 CGATTCACCTGAGCTGACCTGGACGCGAGAGAGCGACTTCCAAATCTACACGGA 780
 OY 781 GGAATCTTTACTGTGAGAGTGTGCTGACGACTTCTTCCGAGGTGACCGGAGATG 840
 Db 781 GGAATCTTTACTGTGAGAGTGTGCTGACGACTTCTTCCGAGGTGACCGGAGATG 840
 OY 841 GGGTGCAGAAATCCAGATTATGTTGCTCAGAAAGAAATCAAGGCGCTGCTCCCTGTT 900
 Db 841 GGGTGCAGAAATCCAGATTATGTTGCTCAGAAAGAAATCAAGGCGCTGCTCCCTGTT 900
 OY 901 CTAAATGCTGACACCACTTACTGTTATGACACCCGGAAATGACTTGGGCAATCACTGA 960
 Db 901 CTAAATGCTGACACCACTTACTGTTATGACACCCGGAAATGACTTGGGCAATCACTGA 960
 OY 961 GTTGTGTGATGCGACAGAGCAATTTGGAGCTGTCTTGAAAGAAACAGATATGATAGT 1020
 Db 961 GTTGTGTGATGCGACAGAGCAATTTGGAGCTGTCTTGAAAGAAACAGATATGATAGT 1020
 OY 1021 TTTTGTACTGTTCTTTCTGATGAGTGTCTGTCGCGCAAGGGCAGTTGATCAAGT 1080
 Db 1021 TTTTGTACTGTTCTTTCTGATGAGTGTCTGTCGCGCAAGGGCAGTTGATCAAGT 1080
 OY 1081 CTCAGAGAGAGCTCTCTGTTTCTTAAGTGTGCTGACAGGGCCACTCTTACTGTAGAAA 1140
 Db 1081 CTCAGAGAGAGCTCTCTGTTTCTTAAGTGTGCTGACAGGGCCACTCTTACTGTAGAAA 1140
 OY 1141 GAGGTACCAAGAAAGCCGCTTATGACAGAGGTGTGAAACAGAGAAATGCAATGT 1200
 Db 1141 GAGGTACCAAGAAAGCCGCTTATGACAGAGGTGTGAAACAGAGAAATGCAATGT 1200
 OY 1201 GGAATGTGAGGTTCCCTTCTCCCTCATGTTCTCATGTTTGTGATGATATTAATCTG 1260
 Db 1201 GGAATGTGAGGTTCCCTTCTCCCTCATGTTCTCATGTTTGTGATGATATTAATCTG 1260
 OY 1261 ATTTACAGACTTAACCTTGTGTGATATTAAGTTTACACGCTGTGTTTATCATCTTTT 1320
 Db 1261 ATTTACAGACTTAACCTTGTGTGATATTAAGTTTACACGCTGTGTTTATCATCTTTT 1320
 OY 1321 GGAAGCCAGAAAGCTTTGGAACGTAATCACTTCCAGATTCGCGATCTCGAC 1380
 Db 1321 GGAAGCCAGAAAGCTTTGGAACGTAATCACTTCCAGATTCGCGATCTCGAC 1380
 OY 1381 TCTTTGCAACAGCACTTCTGCGGAATCTTCTCGGAATGATCACTGACATCCCA 1440
 Db 1381 TCTTTGCAACAGCACTTCTGCGGAATCTTCTCGGAATGATCACTGACATCCCA 1440
 OY 1441 ACCGTGCAACGTGTAACCTTGTGCTTTGCAAAAGAAAGTTGATCTGAATTTCTCTGTGA 1500
 Db 1441 ACCGTGCAACGTGTAACCTTGTGCTTTGCAAAAGAAAGTTGATCTGAATTTCTCTGTGA 1500
 OY 1501 ATTTAGCTTATCAATTCAGAAATAGAGATTGACATGCAACTTTATGAGGTGAGAAA 1560
 Db 1501 ATTTAGCTTATCAATTCAGAAATAGAGATTGACATGCAACTTTATGAGGTGAGAAA 1560
 OY 1561 TTTTATGATGATGTTGGATTCGACCTCATGTTCTGTGTTCTTTATGATGATGTT 1620
 Db 1561 TTTTATGATGATGTTGGATTCGACCTCATGTTCTGTGTTCTTTATGATGATGTT 1620
 OY 1621 TCTATACATGAATCATAGCAAAAATTTTTTGGAAACGTGTGTTGATGATGTTGTTT 1680
 Db 1621 TCTATACATGAATCATAGCAAAAATTTTTTGGAAACGTGTGTTGATGATGTTGTTT 1680

QY 660 ATGAAAAGATGAGAGACCTTAAGCCAAAATTTATCCAGACGAGGCGGAGTACAG 719
 Db 668 ATGAAAAGATGAGAGACCTTAAGCCAAAATTTATCCAGACGAGGCGGAGTACAG 727
 QY 720 CCGATCCACCTCAGCTGAACTGGACGCGACGAGAGCCGATTCMAATCATACTCAGCG 779
 Db 728 CCGATCCACCTCAGCTGAACTGGACGCGACGAGAGCCGATTCMAATCATACTCAGCG 787
 QY 780 AGCAATCTTTTACTGTGAGAGTGGCTGGTCAAGACTTCTTGGAGAGTGGACCGAGATC 839
 Db 788 AGCAATCTTTTACTGTGAGAGTGGCTGGTCAAGACTTCTTGGAGAGTGGACCGAGATC 847
 QY 840 GGGGTGGCAGAAATCCGATTCATGTGTGCTCAGAGAGAAATCAAGCCGTCCTCCCTTGT 899
 Db 848 GGGGTGGCAGAAATCCGATTCATGTGTGCTCAGAGAGAAATCAAGCCGTCCTCCCTTGT 907
 QY 900 TCTAATGCTGCACACCACTTACTGTTCATGGACCCCGGAAATGACTTGGGCAATCAGTCG 959
 Db 908 TCTAATGCTGCACACCACTTACTGTTCATGGACCCCGGAAATGACTTGGGCAATCAGTCG 967
 QY 960 AGTTGTGCTGATCGCACAAGACATTTGGGACTGTCTTGAAGAAAACAGATATGATAGT 1019
 Db 968 AGTTGTGCTGATCGCACAAGACATTTGGGACTGTCTTGAAGAAAACAGATATGATAGT 1027
 QY 1020 GTTTTGTACTGTCTTTCTGAGTACTGTCTGTCTGTCCAGAGGCGAGGCTTATCAGTGA 1079
 Db 1028 GTTTTGTACTGTCTTTCTGAGTACTGTCTGTCTGTCCAGAGGCGAGGCTTATCAGTGA 1087
 QY 1080 GCTCAGAGAGAGACTTCTGTCTTCTAAGTGGCTGCGAGGCGCACTCTTACTGTGATAGA 1139
 Db 1088 GCTCAGAGAGAGACTTCTGTCTTCTAAGTGGCTGCGAGGCGCACTCTTACTGTGATAGA 1147
 QY 1140 AGAGGTACCCACAGAAAGCCGCTAAGTGCAGAGAGTGTGAAAACAGACAGCAATGCAATG 1199
 Db 1148 AGAGGTACCCACAGAAAGCCGCTAAGTGCAGAGAGTGTGAAAACAGACAGCAATGCAATG 1207
 QY 1200 TGGAAATTTGAGCGTTCTTCTTCTCCCTCAGTGTCTCATGTTTGTGCATGATATTACT 1259
 Db 1208 TGGAAATTTGAGCGTTCTTCTTCTCCCTCAGTGTCTCATGTTTGTGCATGATATTACT 1267
 QY 1260 GATTTCACAGACTTACTCTTTGCTGTATATTAAGTATACACCGTGTGTTTATCATCTTT 1319
 Db 1268 GATTTCACAGACTTACTCTTTGCTGTATATTAAGTATACACCGTGTGTTTATCATCTTT 1327
 QY 1320 TGGAAAGCCAGAAAGCGTTTGGAAAACGTATCACCTTCCAGATTTCTGATCTCGA 1379
 Db 1328 TGGAAAGCCAGAAAGCGTTTGGAAAACGTATCACCTTCCAGATTTCTGATCTCGA 1387
 QY 1380 CTCTTTGCAACAGCACTTGTCTGCGAACTCTTCTGGAATGCATTCACACATCCCC 1439
 Db 1388 CTCTTTGCAACAGCACTTGTCTGCGAACTCTTCTGGAATGCATTCACACATCCCC 1447
 QY 1440 AACCGTGCAGGTATCTTGTCTTGTGCAAAAGAGTGTATGTAATTCCTCTGTAG 1499
 Db 1448 AACCGTGCAGGTATCTTGTCTTGTGCAAAAGAGTGTATGTAATTCCTCTGTAG 1507
 QY 1500 AATTAGCTTATACATTTACAGAAATAGCAAGTTCACCTGCCAATCTTTAGTGGTGAGAA 1559
 Db 1508 AATTAGCTTATACATTTACAGAAATAGCAAGTTCACCTGCCAATCTTTAGTGGTGAGAA 1567
 QY 1560 ATTTATAGTTAGTGTGTTGGATGGACCTCAGTTTCTGTGTTTCTTTTATGTGTGTGT 1619
 Db 1568 ATTTATAGTTAGTGTGTTGGATGGACCTCAGTTTCTGTGTTTCTTTTATGTGTGTGT 1627
 QY 1620 TTTCTATACATGATCATATGCCAAAACCTTTTGGAAAACGTGTGGTGTAGATGTGTT 1679
 Db 1628 TTTCTATACATGATCATATGCCAAAACCTTTTGGAAAACGTGTGGTGTAGATGTGTT 1687
 QY 1680 CTTTATACCCCAAGAGATCAAGATATCACTTGTATATTAAGTGTATGATATATTCAT 1739
 Db 1688 CTTTATACCCCAAGAGATCAAGATATCACTTGTATATTAAGTGTATGATATATTCAT 1747
 QY 1740 ACTGTGTATACCTTGGGTGAAAAGTATGCGAGTGGAGACTAAGATGTATTAACCTAAC 1799

Db 1748 ACCGTGTGTACCTTGGGTGAAAAGTATGCGAGTGGAGACTAAGATGTATTAACCTAAC 1807
 QY 1800 TGTGAATCATATGTGTAGAGAAAAGCTGTTCCTCATGTCTTAACAGACTGAATTCMAAGC 1859
 Db 1808 TGTGAATCATATGTGTAGAGAAAAGCTGTTCCTCATGTCTTAACAGACTGAATTCMAAGC 1867
 QY 1860 ATGTCAAGTGAATAGTATGATCTGTGCGATATGAGAGGAGATGCAATGCTTCCCATTC 1919
 Db 1868 ATGTCAAGTGAATAGTATGATCTGTGCGATATGAGAGGAGATGCAATGCTTCCCATTC 1927
 QY 1920 ATTCCTGATGAATTTGTATACATAGTTAACTTGTAAATTTTCTTCTGATGTAATGTG 1979
 Db 1928 ATTCCTGATGAATTTGTATACATAGTTAACTTGTAAATTTTCTTCTGATGTAATGTG 1987
 QY 1980 TATGCTGTGAATATGATATATATTTTGGCTTAAACATACCGTAAACATGTTGTCAAT 2039
 Db 1988 TATGCTGTGAATATGATATATATTTTGGCTTAAACATACCGTAAACATGTTGTCAAT 2047
 QY 2040 TTGAATATCTTAATGCGCAAGTAAACATGCAATGCTTTGGAAATTTGGAAACATGTTTAT 2099
 Db 2048 TTGAATATCTTAATGCGCAAGTAAACATGCAATGCTTTGGAAATTTGGAAACATGTTTAT 2107
 QY 2100 CTTGAGAGCAATATGTTTGTGATTAATGCTTGTATGTTGATATCAAGAAATGAT 2159
 Db 2108 CTTGAGAGCAATATGTTTGTGATTAATGCTTGTATGTTGATATCAAGAAATGAT 2167
 QY 2160 GAACTTCTCAAAACCTGTGTACGTAATGTAAGAGGAGCCGTTTGGAGAGACCA 2219
 Db 2168 GAACTTCTCAAAACCTGTGTACGTAATGTAAGAGGAGCCGTTTGGAGAGACCA 2227
 QY 2220 TTGCATGCTGTCCAAAGTGTCTTGTATAGTGTCTTAACTGSAAGGCTAACCTCAA 2279
 Db 2228 TTGCATGCTGTCCAAAGTGTCTTGTATAGTGTCTTAACTGSAAGGCTAACCTCAA 2287
 QY 2280 AATACCTTTTAACTGCAATCTAATTAATGAGGACAGATATGCTCTTAC 2331
 Db 2288 AATACCTTTTAACTGCAATCTAATTAATGAGGACAGATATGCTCTTAC 2339

RESULT 3
 AAD30156
 ID AAD30156 standard; DNA; 2212 BP.
 XX
 AC AAD30156;
 XX
 DT 17-MAY-2002 (first entry)
 XX
 DE Human MCIP associated DNA #2.
 XX
 XX Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;
 KW heart failure; cardiomyopathy; heart disease; human; gene; ds.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 25..618
 FT /tag= a
 FT /product= "Human MCIP associated protein #2"
 XX
 PN MO200204491-A2.
 XX
 PD 17-JAN-2002.
 XX
 PF 06-JUL-2001; 2001WO-US21662.
 XX
 PR 07-JUL-2001; 2000US-216601P.
 PR 13-FEB-2001; 2001US-0782953.
 XX
 PA (TEXAS) UNIV TEXAS SYSTEM.
 PA (WILL/) WILLIAMS S R.
 PA (ROTH/) ROTHERMEL B.
 XX

QY 1815 GTAGAAAAGCTGTTCCCATGCTCTAAACAGAGACTTGATTCAAAGCATGTCAAGTGATAG 1874
 DB 1836 GTAGAAAAGCTGTTCCCATGCTCTAAACAGAGACTTGATTCAAAGCATGTCAAGTGATAG 1755
 QY 1875 TAGATCTGTGGCGATATGAGAGGATGACGTCCCTTCCCATTCATTCCTGATGGAAT 1934
 DB 1756 TAGATCTGTGGCGATATGAGAGGATGACGTCCCTTCCCATTCATTCCTGATGGAAT 1815
 QY 1935 GTTATACAGAGTAACTTGTATTTTCTAGTGTATGTATGTCTGTAATA 1994
 DB 1816 GTTATACAGAGTAACTTGTATTTTCTAGTGTATGTATGTCTGTAATA 1875
 QY 1995 GGTATATATTTTGGCCTTACATACCCGTACACATGTTTGTCAATTTTGAATACTTAATG 2054
 DB 1876 GGTATATATTTTGGCCTTACATACCCGTACACATGTTTGTCAATTTTGAATACTTAATG 1935
 QY 2055 CCAAGTAAACATGATGCTTTTGAAATTTGAAAGATGCTTTATTTCTTGGAGACCAAT 2114
 DB 1936 CCAAGTAAACATGATGCTTTTGAAATTTGAAAGATGCTTTATTTCTTGGAGACCAAT 1995
 QY 2115 AGCTTTGCAATTAATGCTTTGATGTTCAATCAAGAAATGATGAAAGCTTCTGAAAC 2174
 DB 1996 AGCTTTGCAATTAATGCTTTGATGTTCAATCAAGAAATGATGAAAGCTTCTGAAAC 2055
 QY 2175 CTGTTACCGTACTTGTGTAAGAGGAGCCGCTTTGGAGAGACCATTTGCATGCTGTCCA 2234
 DB 2056 CTGTTACCGTACTTGTGTAAGAGGAGCCGCTTTGGAGAGACCATTTGCATGCTGTCCA 2115
 QY 2235 AGTGTTCCTTTTAAAGTCTTTTAACTGAGAGGCTAACTTCAAAATCTTTTAAAC 2294
 DB 2116 AGTGTTCCTTTTAAAGTCTTTTAACTGAGAGGCTAACTTCAAAATCTTTTAAAC 2175
 QY 2295 TGCATCTATTAATTAATGGGACAGATGCTCTTAC 2331
 DB 2176 TGCATCTATTAATTAATGGGACAGATGCTCTTAC 2212

RESULT 4
 ABN96843
 ID ABN96843 standard; DNA; 2173 BP.
 AC ABN96843;
 XX
 DT 13-AUG-2002 (first entry)
 XX
 DE Gene #3341 used to diagnose liver cancer.
 XX
 KW Gene; liver cancer; db; hepatocellular carcinoma; hepatotropic;
 KW metastatic liver tumour; cytostatic; expression profile; disease state;
 KW disease progression; drug toxicity; drug efficacy; drug metabolism.
 OS Homo sapiens.
 XX
 PN WO200229103-A2.
 XX
 PD 11-APR-2002.
 XX
 PF 02-OCT-2001; 2001WO-US30589.
 XX
 PR 02-OCT-2000; 2000US-237054P.
 XX
 PA (GENE-) GENE LOGIC INC.
 XX
 PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
 DR WPI; 2002-426119/45.
 XX
 PT Diagnosing and detecting the progression of liver cancer,
 PT hepatocellular carcinoma or metastatic liver tumor in a patient,
 PT involves detecting the level of expression of two or more genes in a
 PT liver tissue sample -
 XX
 PS Claim 1; SEQ ID NO 3341; 298bp; English.

XX
 CC The invention relates to a novel method for diagnosing and detecting the
 CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
 CC tumour in a patient, and differentiating metastatic liver cancer from
 CC hepatocellular carcinoma in a patient. Involving detecting the level of
 CC expression of two or more genes represented in ABN93503-ABN97455 in a
 CC tissue sample. The method of the invention has hepatotropic, and
 CC cyostatic activity. The method is useful for diagnosing and detecting
 CC the progression of liver cancer, hepatocellular carcinoma and metastatic
 CC liver carcinoma in a patient. The method is useful for identifying
 CC expression profiles which serve as useful diagnostic markers as well as
 CC markers that can be used to monitor disease states, disease progression,
 CC drug toxicity, drug efficacy and drug metabolism.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WPI
 CC at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).
 XX
 SO Sequence 2173 BP; 606 A; 431 C; 503 G; 633 T; 0 other;
 Query Match 90.1%; Score 2101; DB 24; Length 2173;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 231 GCCAAATTTGAGTCCCTTTAGAGCGTATGACAAAGACATCACTTCAATTTTAAAG 290
 DB 58 GCCAAATTTGAGTCCCTTTAGAGCGTATGACAAAGACATCACTTCAATTTTAAAG 117
 QY 291 AGCTTCAACAGTACAGATTAATTAACCTTCCAGCAACCCCTTCCGAGAGATGCCAGGCTC 350
 DB 118 AGCTTCAACAGTACAGATTAATTAACCTTCCAGCAACCCCTTCCGAGAGATGCCAGGCTC 177
 QY 351 CAGCTGCATTAAGTCACTGATTTCTGGGAAAGAAATGAAGTTATTTTGTCTCAACCTTA 410
 DB 178 CAGCTGCATTAAGTCACTGATTTCTGGGAAAGAAATGAAGTTATTTTGTCTCAACCTTA 237
 QY 411 CACATAGGAAGTCAACACCTGCTCCGCAATCCAGACAGAGCTTCTGATCTCCCT 470
 DB 238 CACATAGGAAGTCAACACCTGCTCCGCAATCCAGACAGAGCTTCTGATCTCCCT 297
 QY 471 CCCGCTCTCCGCGAGTGGATGGAACCAAGTGAAGATGGACCCAGTCATTAATCTAT 530
 DB 298 CCCGCTCTCCGCGAGTGGATGGAACCAAGTGAAGATGGACCCAGTCATTAATCTAT 357
 QY 531 GATCTCTTATATGCTATCTCCAGCTGGGGCCAGGGGAAAGTATGATTCAGACGCG 590
 DB 358 GATCTCTTATATGCTATCTCCAGCTGGGGCCAGGGGAAAGTATGATTCAGACGCG 417
 QY 591 ACTGACACCACTCCGAGGCTGTGTCATGTATGTGAGATGATCAAGAAAGAGGAA 650
 DB 418 ACTGACACCACTCCGAGGCTGTGTCATGTATGTGAGATGATCAAGAAAGAGGAA 477
 QY 651 GAAGAGAAATGAAAGATGAGAGACCTTAAGCCAAAATTTATTCAGACCAAGAGCGCG 710
 DB 478 GAAGAGAAATGAAAGATGAGAGACCTTAAGCCAAAATTTATTCAGACCAAGAGCGCG 537
 QY 711 GAGTACAGCCGATCACTCACTGAGTGAAGTGGCAAGCGAGCGAATCCAAATCAT 770
 DB 538 GAGTACAGCCGATCACTCACTGAGTGAAGTGGCAAGCGAGCGAATCCAAATCAT 597
 QY 771 ACTCAAGGAGAAATCTTTTACTGTGAGAGTGGCTGATCAGACATCTTCCGAGGTCGA 830
 DB 598 ACTCAAGGAGAAATCTTTTACTGTGAGAGTGGCTGATCAGACATCTTCCGAGGTCGA 657
 QY 831 GCCGAGATGGGGTGGCAGAAATCCCAAGTTCATGTTGCTCAGAAAGATCAAGCGCTG 890
 DB 658 GCCGAGATGGGGTGGCAGAAATCCCAAGTTCATGTTGCTCAGAAAGATCAAGCGCTG 717
 QY 891 TCCCTTTGTTTAAAGTGTGACACAGATTAAGTTCATGTCGACCGGGAATGACTTGGGC 950
 DB 718 TCCCTTTGTTTAAAGTGTGACACAGATTAAGTTCATGTCGACCGGGAATGACTTGGGC 777
 QY 951 CAATCACTGAGTTGTGTGATGTCACAAGAGCATTTGGGACTGCTTGAAGAAAACAGAT 1010

Db 778 CAATCACTGAGTTTGTGGTATCGCACAGGACATTTGGGACTGTCTTGAGAAAACAGAT 837
QY 1011 AATGATAGTGTGTTTGTACTGTTGTTCTTTCTGGTAGGTTCTGTCTGTGCAAGGCGAGTT 1070
Db 838 AATGATAGTGTGTTTGTACTGTTGTTCTTTCTGGTAGGTTCTGTCTGTGCAAGGCGAGTT 897
QY 1071 GATCAGTGTAGCTCAGGAGAGACTTCTGTTTCTAAGTGTGGCTGTGAGGGCCACTCTCTA 1130
Db 898 GATCAGTGTAGCTCAGGAGAGACTTCTGTTTCTAAGTGTGGCTGTGAGGGCCACTCTCTA 957
QY 1131 CTGTAGGAAGAGGTACCAAGGAAGCGCTTAGTCAGAGAGGTTGTGAAAACAGGAGC 1190
Db 958 CTGTAGGAAGAGGTACCAAGGAAGCGCTTAGTCAGAGAGGTTGTGAAAACAGGAGC 1017
QY 1191 AATGCAATGTGGAATTTAGAGGTTTCCCTTCTTCCCTCATGTTCTCATGTTTGTGCAATG 1250
Db 1018 AATGCAATGTGGAATTTAGAGGTTTCCCTTCTTCCCTCATGTTCTCATGTTTGTGCAATG 1077
QY 1251 TATATTACTGATTTACAGAGACTAACCTTTGTTTCGTATATATAAGTTACACCGTTGTTGTTT 1310
Db 1078 TATATTACTGATTTACAGAGACTAACCTTTGTTTCGTATATATAAGTTACACCGTTGTTGTTT 1137
QY 1311 TACATCTTTTGGGAAGCCAGGAAGCGTTTGGAAAAAGTATACCTTTTCCAGATTTCTCG 1370
Db 1138 TACATCTTTTGGGAAGCCAGGAAGCGTTTGGAAAAAGTATACCTTTTCCAGATTTCTCG 1197
QY 1371 GATTCGACTCTTTGCAAGAGACTTGTCTTGGGAACTCTTCTCGGAATGCATTCACCTC 1430
Db 1198 GATTCGACTCTTTGCAAGAGCACTTGTCTTGGGAACTCTTCTCGGAATGCATTCACCTC 1257
QY 1431 AGCATCCCAACCGTGCAACGTTGTAACCTTGTGCTTTTGCAAAAGAAAGTTGATCGAAAT 1490
Db 1258 AGCATCCCAACCGTGCAACGTTGTAACCTTGTGCTTTTGCAAAAGAAAGTTGATCGAAAT 1317
QY 1491 CCTCTGAGAAATTTAGCTTATACAAATTCAGAGAATAGCAGTTTCACTGCCAACTTTTAGT 1550
Db 1318 CCTCTGAGAAATTTAGCTTATACAAATTCAGAGAATAGCAGTTTCACTGCCAACTTTTAGT 1377
QY 1551 GGGTGAGAAATTTAGTTTGTAGTTGTTGGGATCGGACCTCAGTTTCGTGTTTCTTTTA 1610
Db 1378 GGGTGAGAAATTTAGTTTGTAGTTGTTGGGATCGGACCTCAGTTTCGTGTTTCTTTTA 1437
QY 1611 TGTGTTGGTTTCTATACATGAATCATAGCCAAAACCTTTTTCGAACTGTTGTTGAG 1670
Db 1438 TGTGTTGGTTTCTATACATGAATCATAGCCAAAACCTTTTTCGAACTGTTGTTGAG 1497
QY 1671 TAGTTGTTCTTTTACCCCAAGAGACATCAAGATACACTTGTAAATAAAGCTGATAGCA 1730
Db 1498 TAGTTGTTCTTTTACCCCAAGAGACATCAAGATACACTTGTAAATAAAGCTGATAGCA 1557
QY 1731 TATATTACATCTGTTGTACACTTGGGTGAAAGTATGCGAGTGGGAGACTAAGATGTAT 1790
Db 1558 TATATTACATCTGTTGTACACTTGGGTGAAAGTATGCGAGTGGGAGACTAAGATGTAT 1617
QY 1791 TAACTCACTGTGGAATCATATGTTGTAGAAAAGCTGTTCCCATGTCTAAACAGGACTTGA 1850
Db 1618 TAACTCACTGTGGAATCATATGTTGTAGAAAAGCTGTTCCCATGTCTAAACAGGACTTGA 1677
QY 1851 ATTCAAGCATGTCAAGTGTAGTATGCTGTGGCGATATGAGAGGGATGCAAGTGCCTT 1910
Db 1678 ATTCAAGCATGTCAAGTGTAGTATGCTGTGGCGATATGAGAGGGATGCAAGTGCCTT 1737
QY 1911 TCCCATTCATCTCGATGGAATTTGTATAGTTTACATTTGTAATTTTTTCTAGT 1970
Db 1738 TCCCATTCATCTCGATGGAATTTGTATAGTTTACATTTGTAATTTTTTCTAGT 1797
QY 1971 TGTAAATGTGTATGCTGTGTAATAGGTATATATATTTTGGCCCTTACAATACCGTAAACAATG 2030
Db 1798 TGTAAATGTGTATGCTGTGTAATAGGTATATATATTTTGGCCCTTACAATACCGTAAACAATG 1857
QY 2031 TTGTGCTATTTGGAATTAATGCAAGTAAACAATGCAATGCAATGCTTTTGGAAATTTGGAAGAT 2090
Db 1858 TTGTGCTATTTGGAATTAATGCAAGTAAACAATGCAATGCAATGCTTTTGGAAATTTGGAAGAT 1917

QY 2091 GGTATTATTTCTTTGAGAAGCAAAATATGTTTGCATTAAATGCTTTGATTGTTTCATATCAAG 2150
Db 1918 GGTATTATTTCTTTGAGAAGCAAAATATGTTTGCATTAAATGCTTTGATTGTTTCATATCAAG 1977
QY 2151 AATTTGATTGAAGCTTCTCAAAACCCCTGTTTACGGTACTTGGTAAAGGGAGCGGTTTGG 2210
Db 1978 AATTTGATTGAAGCTTCTCAAAACCCCTGTTTACGGTACTTGGTAAAGGGAGCGGTTTGG 2037
QY 2211 GAGAGACCAATTCGATCGCTGTCCAAAGTGTTCCTGTTAAGTCTCTTTAAACTGGAGAGC 2270
Db 2038 GAGAGACCAATTCGATCGCTGTCCAAAGTGTTCCTGTTAAGTCTCTTTAAACTGGAGAGC 2097
QY 2271 TAACTTCAAAATACCTTTTAACTGCATCTATATAAATGGGCACAGTATGCTCTTFA 2330
Db 2098 TAACTTCAAAATACCTTTTAACTGCATCTATATAAATGGGCACAGTATGCTCTTFA 2157
QY 2331 C 2331
Db 2158 C 2158

RESULT 5
AAF18328
ID AAF18328 standard; DNA; 2358 BP.
XX AC AAF18328;
XX XX
DT 14-MAR-2001 (first entry)
XX XX
XX Lung cancer associated polynucleotide sequence SEQ ID 347.
XX Human; lung cancer associated protein; neuroprotective; cytostatic;
KW cardioactive; immunomodulatory; muscular active; vulnerary;
KW gastrointestinal; nephrotropic; antiinfective; gynecological;
KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
KW proliferative disorder; wound healing; infectious disease; ds.
XX OS Homo sapiens.
XX OS
PN WO200055180-A2.
XX XX
PD 21-SEP-2000.
XX XX
PF 08-MAR-2000; 2000WO-US05918.
XX XX
PR 12-MAR-1999; 99US-0124270.
XX XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (ROSE/) ROSEN C A.
XX XX
PI Ruben SM;
XX XX
DR WPI; 2000-587514/55.
DR P-PSDB; AAB58452.
XX XX
PT Lung cancer associated gene sequences, referred to as lung cancer
PT antigens, useful for treatment, prevention, and diagnosis of disorders
PT such as lung cancer -
XX XX
PS Claim 1; Page 805-806; 1425pp; English.
XX XX
CC Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer
CC associated proteins represented in AAB58106 - AAB58548. Lung cancer
CC associated proteins and polynucleotide sequences, their agonists, and
CC antagonists may have neuroprotective; cytostatic; cardioactive;
CC immunomodulatory; muscular active general; vulnerary; gastrointestinal
CC general; nephrotropic; antiinfective; gynecological; or antibacterial
CC activity. The invention also includes antibodies specific for the
CC protein or polynucleotide sequences. The lung cancer associated
CC polynucleotide sequences may be used for detection of lung cancer,
CC chromosome identification, as chromosome markers, and for numerous other
CC diagnostic or research purposes. The proteins may be used to treat

disorders such as neural, immune, muscular, reproductive, gastrointestinal, pulmonary, cardiovascular, renal, and proliferative disorders. The proteins may also be used in the treatment of wounds and infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and peptide AAB58549 are used in the course of the invention for the identification and characterisation of the polynucleotide and protein sequences.

Sequence 2358 BP; 641 A; 482 C; 584 G; 647 T; 4 other:

Query Match 89.8%; Score 2094.2; DB 21; Length 2358;
Best Local Similarity 97.3%; Pred. No. 0;
Matches 2137; Conservative 3; Mismatches 56; Indels 1; Gaps 1;

135 GAAAGCAGATGATTTTAACTTAACTACAGTTTACCTCCCTGATTCGCTGTG 194
105 GACTGCGAGATGAGAGGTGAGCTGCGAGACCTGCCAGGCGCAATCGCTGTAC 164
195 GCAAGCAGATGATTTTCAAGCGAAAGTGAACCGAGGCGCAATTTGAGTCCCTTTAGG 254
165 CTGAGACCCCGCGCTGTTCTGAGCGGCTGTGCGGCGCAATTTGAGTCCCTTTAGG 224
255 ACCTATGAGAGGACATCACTTTAGATTTTAAAGCTTCAACGAGTCAAGATAAC 314
225 ACCTATGAGAGGACATCACTTTAGATTTTAAAGCTTCAACGAGTCAAGATAAC 284
315 TTGAGCAGACCTTCTCCGAGAGATGCCAGGCTCAGTGTGATGAGCTGAGTTCTG 374
285 TTGAGCAGACCTTCTCCGAGAGATGCCAGGCTCAGTGTGATGAGCTGAGTTCTG 344
375 GGAAGGAAATGAAGTTATTTTGTCTCAGACCTTACATAGAGAGCTCACCTGCT 434
345 GGAAGGAAATGAAGTTATTTTGTCTCAGACCTTACATAGAGAGCTCACCTGCT 404
435 CGCCCAATTCAGACAGACGATTTCTGATCTCCCTCCGCTCTCCGAGTGGATGG 494
405 CGC-CAATTCAGACAGACGATTTCTGATCTCCCTCCGCTCTCCGAGTGGATGG 463
495 AATCAATGAGAGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 554
464 AATCAATGAGAGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 523
555 CTGGGCGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 614
524 CTGGGCGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 583
615 GTCCATGATGTGAGAGTGAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGG 674
584 GTCCATGATGTGAGAGTGAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGG 643
675 AGACCTAAGCCAAATTTATCAAGACAGAGAGGCGGAGTACAGCCGATCCAGC 734
644 AGACCTAAGCCAAATTTATCAAGACAGAGAGGCGGAGTACAGCCGATCCAGC 703
735 TGAATCGGACGCGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 794
704 TGAATCGGACGCGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 763
795 TGAAGTGTGCTGTACAGACTTTCTCGAGGTGAGCGGAGTCCGGGAGGAGAGATC 854
764 TGAAGTGTGCTGTACAGACTTTCTCGAGGTGAGCGGAGTCCGGGAGGAGAGATC 823
855 CCAAGTTCATGTTGCTCAGAGAGAGATCAAGGCGTGTCCCTTTCTAATGTGCAAC 914
824 CCAAGTTCATGTTGCTCAGAGAGAGATCAAGGCGTGTCCCTTTCTAATGTGCAAC 883
915 CAGTTCATGTTGCTCAGAGAGAGATCAAGGCGTGTCCCTTTCTAATGTGCAAC 974
884 CAGTTCATGTTGCTCAGAGAGAGATCAAGGCGTGTCCCTTTCTAATGTGCAAC 943
975 CACAAGGAGATTTGGAGAGTGTGAGAGAGAGATTAAGTGTGTTTGTACTGTTTC 1034
944 CACAAGGAGATTTGGAGAGTGTGAGAGAGAGATTAAGTGTGTTTGTACTGTTTC 1003

1035 TTTCTGTAGTGTCTGTCTGTGTCAGAGGAGAGTGTGATCAGTACCTCAGAGAGAGCT 1094
1004 TTTCTGTAGTGTCTGTCTGTGTCAGAGGAGAGTGTGATCAGTACCTCAGAGAGAGCT 1063
1095 TCTGTGTCTAAGTGTGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1154
1064 TCTGTGTCTAAGTGTGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1123
1155 AGCGGCTAGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1214
1124 AGCGGCTAGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1183
1215 TTTCTGTCTTCCCTCATGTTCTCATGTTGTGATGATTTTACTGATTTTCAAGACTAA 1274
1184 TTTCTGTCTTCCCTCATGTTCTCATGTTGTGATGATTTTACTGATTTTCAAGACTAA 1243
1275 CTTTGTGTCTAATTAAGTGAACCGTGTGTTTGTGATGATTTTACTGATTTTCAAGACTAA 1334
1244 CTTTGTGTCTAATTAAGTGAACCGTGTGTTTGTGATGATTTTACTGATTTTCAAGACTAA 1303
1335 GCGTTTGTGAGAGGAGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1394
1304 GCGTTTGTGAGAGGAGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1363
1395 CTGCTGTGAGAGGAGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1454
1364 CTGCTGTGAGAGGAGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1423
1455 AACTTGTGCTTTTCAAGAGAGGAGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1514
1424 AACTTGTGCTTTTCAAGAGAGGAGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1483
1515 ATTGAGAGATGAGAGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1574
1484 ATTGAGAGATGAGAGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1543
1575 TTTGGAGTGTGAGAGGAGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1634
1544 TTTGGAGTGTGAGAGGAGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1603
1635 ATAGCAAAAGCTTTTGTGAGAGGAGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1694
1604 ATAGCAAAAGCTTTTGTGAGAGGAGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1663
1695 GACATCAAGATACACTTGTAAATTAAGCTGATGATGATGATGATGATGATGATGATG 1754
1664 GACATCAAGATACACTTGTAAATTAAGCTGATGATGATGATGATGATGATGATGATG 1723
1755 GGGTGAAGAGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1814
1724 GGGTGAAGAGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1783
1815 GTAGGAGAGAGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1874
1784 GTAGGAGAGAGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1843
1875 TTAGATGTGCGAGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1934
1844 TTAGATGTGCGAGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1903
1935 GTTATCTAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1994
1904 GTTATCTAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1963
1995 GGTATTAATTTTGGCTTACAAATCCGTAACAGTGTGTGATTTGAAATCTTAATG 2054
1964 GGTATTAATTTTGGCTTACAAATCCGTAACAGTGTGTGATTTGAAATCTTAATG 2023
2055 CCAAGTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2114
2024 CCAAGTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2083

QY 2115 ATGTTTGCATTAATAGCTTTGATGTTTCATATCAAGAAATGATGAACGTTCTCAAAACC 2174
 Db |||||||
 QY 2084 ATGTTTGCATTAATAGCTTTGATGTTTCATATCAAGAAATGATGAACGTTCTCAAAACC 2143
 Db |||||||
 QY 2175 CTGTTTACGGTACTTGGTAAAGAGGAGCGGTTTGGGAGAGACCAATTCGATCGCTGTCCA 2234
 Db |||||||
 QY 2144 CTGTTTACGGTACTTGGTAAAGAGGAGCGGTTTGGGAGAGACCAATTCGATCGCTGTCCA 2203
 Db |||||||
 QY 2235 AGTGTTCCTGTTAGTGTCTTTTAACTGGAGAGGCTTAACCTCAAAATACCTTTTAAAC 2294
 Db |||||||
 QY 2204 AGTGTTCCTGTTAGTGTCTTTTAACTGGAGAGGCTTAACCTCAAAATACCTTTTAAAC 2263
 Db |||||||
 QY 2295 TGCATTCCTATAATAAATGGGACACGATGCTCCCTTAC 2331
 Db |||||||
 QY 2264 TGCATTCCTATAATAAATGGGACACGATGCTCCCTTAC 2300
 Db |||||||

RESULT 6
 AAX01282
 ID AAX01282 standard; cDNA to mRNA; 2174 BP.
 XX
 AC AAX01282;
 XX
 DT 09-APR-1999 (first entry)
 XX
 DE Human DSCR1 coding sequence.
 XX
 KW DSCR1; Down's Syndrome critical region 1; human; pathogenic abnormality;
 KW Central Nervous System development; mental retardation; heart defect; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 49..564
 FT /*tag= a
 FT polyA_signal 1541..1546
 FT /*tag= b
 FT polyA_signal 2132..2137
 FT /*tag= c
 XX
 US5869318-A.
 XX
 XX 09-FEB-1999.
 PD
 XX 07-JUN-1996; 96US-0665040.
 XX
 XX 07-JUN-1995; 95BS-0001140.
 XX
 PA (PALL/) PALLEJA X E.
 XX
 XX Puentes JJ, Palleja XE, Pritchard M;
 XX
 WPI: 1999-152781/13.
 DR P-PSDB; AA73898.
 XX
 XX DNA encoding foetal brain proteins - believed to be associated with
 PT Down's syndrome
 XX
 XX Claim 1; Column 15-18; 19pp; English.
 CC
 CC This sequence is the Down's Syndrome critical region 1 (DSCR1) gene
 CC of the invention. The DSCR1 gene was found to be located in the
 CC q22.1-22.2 region of human chromosome 21. An increase in the transient
 CC expression of DSCR1 mRNA in the brains of young rats, compared to
 CC expression levels in the brains of adult rats, suggests an important role
 CC for DSCR1 during the development of the Central Nervous System (CNS), and
 CC that over expression of DSCR1 may be involved in pathogenic abnormalities
 CC of mental retardation and/or heart defects as found in Down's syndrome
 CC patients.
 XX
 SQ Sequence 2174 BP; 605 A; 436 C; 502 G; 528 T; 3 other;
 Query Match 88.3%; Score 2059.2; DB 20; Length 2174;

Best Local Similarity 99.0%; Pred. No. 0;
 Matches 2081; Conservative 0; Mismatches 20; Indels 1; Gaps 1;

QY 231 GCCAAATTTGAGTCCCTCTTTAGGACGTATGACAGGACATCACCTTTTCAGTATTTTAAAG 290
 Db |||||||
 QY 291 AGCTTCAGGAGTCAAGATAAATTTTCAGGAAACCCCTTCTCCGACGAGATGCCAGGCTC 350
 Db |||||||
 QY 118 AGCTTCAGGAGTCAAGATAAATTTTCAGGAAACCCCTTCTCCGACGAGATGCCAGGCTC 177
 Db |||||||
 QY 351 CAGCTGCATAAGACTGAGTTTCTGGGAAAGGAAATGAAGTTATATTTTCTCAGACCTTA 410
 Db |||||||
 QY 178 CAGCTGCATAAGACTGAGTTTCTGGGAAAGGAAATGAAGTTATATTTTCTCAGACCTTA 237
 Db |||||||
 QY 411 CACATAGGAACTCACACCTGGCTCCGCAAAATCCAGACAAGCAGTTTCTGATCTCCCT 470
 Db |||||||
 QY 238 CACATAGGAACTCACACCTGGCTCCGCAAAATCCAGACAAGCAGTTTCTGATCTCCCT 297
 Db |||||||
 QY 471 CCGGCTCTCCGCGCAGTGGGATGGAACAAGTGGAGAGATGGACCCAGTCATAAATAT 530
 Db |||||||
 QY 298 CCGGCTCTCCGCGCAGTGGGATGGAACAAGTGGAGAGATGGACCCAGTCATAAATAT 357
 Db |||||||
 QY 531 GATCTTTATATGCCATCTCCAAAGCTGGGGCCAGGGGAAAGTATGAATTCGACGCGG 590
 Db |||||||
 QY 358 GATCTTTATATGCCATCTCCAAAGCTGGGGCCAGGGGAAAGTATGAATTCGACGCGG 417
 Db |||||||
 QY 591 ACTGACACACTCCGAGCGTGGTGCATGATGTGAGAGTGTATCAAGAGAAGAGGAA 650
 Db |||||||
 QY 418 ACTGACACACTCCGAGCGTGGTGCATGATGTGAGAGTGTATCAAGAGAAGAGGAA 477
 Db |||||||
 QY 651 GAAAGAGGAAATGGAAGAAATGAGGAGACCTAAGCCAAAATTTCCAGACCAGGAGCGG 710
 Db |||||||
 QY 478 GAAGAGGAAATGGAAGAAATGAGGAGACCTAAGCCAAAATTTCCAGACCAGGAGCGG 537
 Db |||||||
 QY 711 GAGTACACGCCGATCCACTCAGCTGAACTGGCACGGCAGCAGCAGCATTCCAAATCAT 770
 Db |||||||
 QY 538 GAGTACACGCCGATCCACTCAGCTGAACTGGCACGGCAGCAGCAGCATTCCAAATCAT 597
 Db |||||||
 QY 771 ACTCAGCGGAGGAAATCTTTTACTGTGAGAGTGGTGTGACAGCTTCTTCGAGAGTGCA 830
 Db |||||||
 QY 598 ACTCAGCGGAGGAAATCTTTTACTGTGAGAGTGGTGTGACAGCTTCTTCGAGAGTGCA 657
 Db |||||||
 QY 831 GCCAGATTCGGGTTGGCAGAAATCCCAAGTTCATGCTGCTCAGAGAAGAAATCAAGGCCGTG 890
 Db |||||||
 QY 658 GCCAGATTCGGGTTGGCAGAAATCCCAAGTTCATGCTGCTCAGAGAAGAAATCAAGGCCGTG 717
 Db |||||||
 QY 891 TCCCTTGTCTTAATGTGCAACCACTTACTGTTTCAATGGCACCCCGGGAATGACTTGGGC 950
 Db |||||||
 QY 718 TCCCTTGTCTTAATGTGCAACCACTTACTGTTTCAATGGCACCCCGGGAATGACTTGGGC 777
 Db |||||||
 QY 951 CAATCACTGAGTTTGTGTTGATCGCAACAGGACATTTGGGACTGTCTTGAGAAAACAGAT 1010
 Db |||||||
 QY 778 CAATCACTGAGTTTGTGTTGATCGCAACAGGACATTTGGGACTGTCTTGAGAAAACAGAT 837
 Db |||||||
 QY 1011 AATGATAGTGTTTGTACTTTGTTCTTTTCTGAGAGTGTCTGTCGCCAAGGCGAGTT 1070
 Db |||||||
 QY 838 AATGATAGTGTTTGTACTTTGTTCTTTTCTGAGAGTGTCTGTCGCCAAGGCGAGTT 897
 Db |||||||
 QY 1071 GATCAGTGAGCTCAGGAGAGAGCTTCTGTTTCTAAGTGGCCTGCAGGGGCCACTCTCTA 1130
 Db |||||||
 QY 898 GATCAGTGAGCTCAGGAGAGAGCTTCTGTTTCTAAGTGGCCTGCAGGGGCCACTCTCTA 957
 Db |||||||
 QY 1131 CTGTTAGGAGAGGTACCAACAGAGCCGCTAGTCGACAGAGGTTGTGAAAACAGCAGC 1190
 Db |||||||
 QY 958 CTGTTAGGAGAGGTACCAACAGAGCCGCTAGTCGACAGAGGTTGTGAAAACAGCAGC 1017
 Db |||||||
 QY 1191 AATGCAATGTGGAATTTGTAGCGTTTCTTTTCTTCCCTCATGTTCTCATGTTTGTGCATG 1250
 Db |||||||
 QY 1018 AATGCAATGTGGAATCTGTAGCGTTTCTTTTCTTCCCTCATGTTCTCATGTTTGTGCATG 1077
 Db |||||||
 QY 1251 TATATTACTGATTTACAAGACTAACCTTTGTCGTATATAAAGTTACACCGTTGTGTTT 1310
 Db |||||||

04-FEB-2000; 2000US-0180312.
 26-MAY-2000; 2000US-0207456.
 30-JUN-2000; 2000US-0608408.
 03-AUG-2000; 2000US-0632366.
 21-SEP-2000; 2000US-0234687.
 27-SEP-2000; 2000US-0236359.
 04-OCT-2000; 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.

Penn SG, Hanzel DK, Chen W, Rank DR;
 WPI, 2001-48899/53.

Single exon nucleic acid probes for analyzing gene expression in human hearts -

Claim 1; SEQ ID No 809; 530bp; English.

The present invention relates to single exon nucleic acid probes for measuring human gene expression in a sample derived from human heart. The present sequence is one such probe. The probes may be used for predicting, measuring and displaying gene expression in samples derived from the human heart via microarrays. By measuring gene expression, the probes are useful for predicting, diagnosing, grading, staging, monitoring and prognosing diseases of the human heart and vascular system e.g. cardiovascular disease, hypertension, cardiac arrhythmias and congenital heart disease.
 Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WPI at ftp.wpi.int/pub/published_pct_sequences.

Sequence 486 BP; 113 A; 139 C; 104 G; 130 T; 0 other;

Query Match 20.8%; Score 486; DB 22; Length 486;
 Best Local Similarity 100.0%; Pred. No. 1.8e-129; Indels 0; Gaps 0;
 Matches 486; Conservative 0; Mismatches 0;

570 AAGTATGAATTGCGACGAGGAGTACACCACTCCGCGTGTGTCTATGTATGTAG 629
 486 AAGTATGAATTGCGACGAGGAGTACACCACTCCGCGTGTGTCTATGTATGTAG 427
 630 AGTATCAAG 689
 426 AGTATCAAG 367
 690 ATTATCCAGACGAGAGGCGCGAGTACACCGCGATCCAGCTGAGCTGAGCGAG 749
 366 ATTATCCAGACGAGAGGCGCGAGTACACCGCGATCCAGCTGAGCTGAGCGAG 307
 750 CGAGAGCGCATTCGCAATCATCTACTCAGCGAGAGATCTTTTACTGTGAGGTGGT 809
 306 CGAGAGCGCATTCGCAATCATCTACTCAGCGAGAGATCTTTTACTGTGAGGTGGT 247
 810 ACGACTTCTTGGAGGTGGACCGAGATCGGGTGGCAGAAATCCAGTTCACTTGGT 869
 246 ACGACTTCTTGGAGGTGGACCGAGATCGGGTGGCAGAAATCCAGTTCACTTGGT 187
 870 CAGAGAGATTCAGGCGCGTGTCCCTGTTCTAATGCTGACACAGTACTTGTGATG 929
 186 CAGAGAGATTCAGGCGCGTGTCCCTGTTCTAATGCTGACACAGTACTTGTGATG 127
 930 GCACCCGGGAAATGACTTGGGCAATCACTGAGTTGTGTGATGTCACAGAGACATT 989
 126 GCACCCGGGAAATGACTTGGGCAATCACTGAGTTGTGTGATGTCACAGAGACATT 67
 990 GACTCTTTGAGAGAGAGAGATATATAGTGTGTGTGATGTTCTTTTCTGGTAGGTT 1049
 66 GACTCTTTGAGAGAGAGAGATATATAGTGTGTGTGATGTTCTTTTCTGGTAGGTT 7
 1050 TGTCTG 1055
 6 TGTCTG 1

RESULT 10

AAK00816/C
 ID AAK00816 standard; DNA; 486 BP.

AAK00816;

05-NOV-2001 (first entry)

Human brain expressed single exon probe SEQ ID NO: 807.

Human; brain expressed exon; gene expression analysis; probe;

microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;

epilepsy; cancer; ss.

Homo sapiens.

W0200157275-A2.

09-AUG-2001.

30-JAN-2001; 2001MO-US00667.

04-FEB-2000; 2000US-0180312.

26-MAY-2000; 2000US-0207456.

30-JUN-2000; 2000US-0608408.

03-AUG-2000; 2000US-0632366.

21-SEP-2000; 2000US-0234687.

27-SEP-2000; 2000US-0236359.

04-OCT-2000; 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.

Penn SG, Hanzel DK, Chen W, Rank DR;

WPI, 2001-483446/52.

Single exon nucleic acid probes for analyzing gene expression in human brains -

Example 4; SEQ ID NO: 807; 650bp + Sequence listing; English.

The present invention provides a number of single exon nucleic acid probes which are derived from genomic sequences expressed in human brain. They can be used to measure gene expression in brain cell samples, which may enable the diagnosis and improved treatment of nervous system diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia, epilepsy and cancers. The present sequence is one of the probes of the invention.

Sequence 486 BP; 113 A; 139 C; 104 G; 130 T; 0 other;

Query Match 20.8%; Score 486; DB 22; Length 486;
 Best Local Similarity 100.0%; Pred. No. 1.8e-129; Indels 0; Gaps 0;
 Matches 486; Conservative 0; Mismatches 0;

570 AAGTATGAATTGCGACGAGGAGTACACCACTCCGCGTGTGTCTATGTATGTAG 629
 486 AAGTATGAATTGCGACGAGGAGTACACCACTCCGCGTGTGTCTATGTATGTAG 427
 630 AGTATCAAG 689
 426 AGTATCAAG 367
 690 ATTATCCAGACGAGAGGCGCGAGTACACCGCGATCCAGCTGAGCTGAGCGAG 749
 366 ATTATCCAGACGAGAGGCGCGAGTACACCGCGATCCAGCTGAGCTGAGCGAG 307
 750 CGAGAGCGCATTCGCAATCATCTACTCAGCGAGAGATCTTTTACTGTGAGGTGGT 809
 306 CGAGAGCGCATTCGCAATCATCTACTCAGCGAGAGATCTTTTACTGTGAGGTGGT 247

QY	81C	ACGACTTCTTCGGAGTGGCAGCCGAGATCGGGGTGGCAGAAATCCAGTTCATCTTGCT	869			
Db	246	ACGACTTCTTCGGAGTGGCAGCCGAGATCGGGGTGGCAGAAATCCAGTTCATCTTGCT	187			
QY	870	CAGAAAGAAATCAAGGCGGTGTCCTTCTTAATGCTGCACACAGTTACTGTTTCATG	929			
Db	186	CAGAAAGAAATCAAGGCGGTGTCCTTCTTAATGCTGCACACAGTTACTGTTTCATG	127			
QY	930	GCACCCGGGAATGACTTGGGCCAATCACTGAGTTTGTGGTGATCGCACAGGACATTTGG	989			
Db	126	GCACCCGGGAATGACTTGGGCCAATCACTGAGTTTGTGGTGATCGCACAGGACATTTGG	67			
QY	990	GACTGTCTTGGAGAAAACAGATAATGATAGTGTGTTTGTACTGTTCTTTCTGGTAGGTTTC	1049			
Db	66	GACTGTCTTGGAGAAAACAGATAATGATAGTGTGTTTGTACTGTTCTTTCTGGTAGGTTTC	7			
QY	1050	TGTCGTG 1055				
Db	6	TGTCGTG 1				
RESULT 11						
AAK26270/c						
ID	AAK26270 standard; DNA; 486 BP.					
XX						
AC	AAK26270;					
XX						
DT	06-NOV-2001 (first entry)					
XX						
DE	Human bone marrow expressed single exon probe SEQ ID NO: 827.					
XX						
KW	Human; bone marrow expressed exon; gene expression analysis; probe;					
KW	microarray; cancer; leukaemia; lymphoma; myeloma; ss.					
XX						
OS	Homo sapiens.					
XX						
PN	W0200157276-A2.					
XX						
PD	09-AUG-2001.					
XX						
PF	30-JAN-2001; 2001WO-US00668.					
XX						
PR	04-FEB-2000; 2000US-0180312.					
PR	26-MAY-2000; 2000US-0207456.					
PR	30-JUN-2000; 2000US-0608408.					
PR	03-AUG-2000; 2000US-0632366.					
PR	21-SEP-2000; 2000US-0234687.					
PR	27-SEP-2000; 2000US-0236359.					
PR	04-OCT-2000; 2000GB-0024263.					
XX						
PA	(MOLE-) MOLECULAR DYNAMICS INC.					
XX						
PI	Penn SG, Hanzel DK, Chen W, Rank DR;					
XX						
DR	WPI; 2001-488900/53.					
XX						
PT	Human genome-derived single exon nucleic acid probes useful for					
PT	analyzing gene expression in human bone marrow -					
XX						
PS	Example 4; SEQ ID NO: 827; 658pp + Sequence Listing; English.					
XX						
CC	The present invention provides a number of single exon nucleic acid					
CC	probes which are derived from genomic sequences expressed in the human					
CC	bone marrow. They can be used to measure gene expression in bone marrow					
CC	samples, which may enable the improved diagnosis and treatment of cancer					
CC	such as lymphoma, leukaemia and myeloma. The present sequence is one of					
CC	the probes of the invention.					
XX						
SQ	Sequence 486 BP; 113 A; 139 C; 104 G; 130 T; 0 other;					
Query Match						
Best Local Similarity 20.8%; Score 486; DB 22; Length 486;						
Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps						

PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human cervical epithelial cells -

PS Claim 25; SEQ ID No 836; 487bp; English.

XX The present invention relates to human single exon nucleic acid probes
 CC (SENPs). The present sequence is one such probe. The SENPs are derived
 CC from human HeLa cells. The SENPs can be used to produce a single exon
 CC microarray, which can be used for measuring human gene expression in a
 CC sample derived from human cervical epithelial cells. By measuring gene
 CC expression, the probes are therefore useful in grading and/or staging
 CC of diseases of the cervix, notably cervical cancer.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 486 BP; 113 A; 139 C; 104 G; 130 T; 0 other;

Query Match 20.8%; Score 486; DB 22; Length 486;
 Best Local Similarity 100.0%; Pred. No. 1.8e-129; Indels 0; Gaps 0;
 Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 570 AAGTATGAAATTCAGCGCAGCAGTGCACACCTCCCGCGTGTGCTCATGTATGTAG 629
 DB 486 AAGTATGAAATTCAGCGCAGCAGTGCACACCTCCCGCGTGTGCTCATGTATGTAG 427
 QY 630 AGTATCAAG 689
 DB 426 AGTATCAAG 367
 QY 690 ATTATCCAGACCGAGCGCGAGTACAGCGCGATCCAGCTCAGCTGAACTGGCAGCGGA 749
 DB 366 ATTATCCAGACCGAGCGCGAGTACAGCGCGATCCAGCTCAGCTGAACTGGCAGCGGA 307
 QY 750 CGAGAGCGCATTCGAAATCATCTACCGGAGAGATCTTTTACTGTGAGGTGCTGCTC 809
 DB 306 CGAGAGCGCATTCGAAATCATCTACCGGAGAGATCTTTTACTGTGAGGTGCTGCTC 247
 QY 810 ACGACTCTTTCGAGGTGCGACCGCGAGATCGGGGTGGCAGAAATCCCACTTCTGCT 869
 DB 246 ACGACTCTTTCGAGGTGCGACCGCGAGATCGGGGTGGCAGAAATCCCACTTCTGCT 187
 QY 870 CAGAGAGAAATCAAGCGCGTGTCCCTTGTCTAATGCTGCACACCACTTACTGTCATG 929
 DB 186 CAGAGAGAAATCAAGCGCGTGTCCCTTGTCTAATGCTGCACACCACTTACTGTCATG 127
 QY 930 GCACCCGGGAATGACTTGGGCAATCACTGACTTGTGTGATTCGCAAGAGACATTTGG 989
 DB 126 GCACCCGGGAATGACTTGGGCAATCACTGACTTGTGTGATTCGCAAGAGACATTTGG 67
 QY 990 GACTGTCTTGAGAAACAGATATGATGATGTTTGTACTTGTCTTCTTCTGAGTTTC 1049
 DB 66 GACTGTCTTGAGAAACAGATATGATGATGTTTGTACTTGTCTTCTTCTGAGTTTC 7
 QY 1050 TGTCTG 1055
 DB 6 TGTCTG 1

RESULT 13

AA132163/c
 ID AA132163 standard; DNA; 486 BP.

AC AA132163;

XX 17-OCT-2001 (first entry)

XX Probe #849 used to measure gene expression in human placenta sample.

XX Probe; microarray; human; placenta; antenatal diagnosis;

KM genetic disorder; ss.

XX Homo sapiens.

OS

XX PN WO200157272-A2.
 XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US000663.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0633366.

XX PR 21-SEP-2000; 2000US-0234687.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-48897/53.

XX Human genome-derived single exon nucleic acid probes useful for

XX analyzing gene expression in human placenta -

XX Claim 25; SEQ ID No 849; 654bp; English.

XX The present invention relates to single exon nucleic acid probes (SENPs).

XX The present sequence is one such probe. The probes are useful for

XX producing a microarray for predicting, measuring and displaying gene

XX expression in samples derived from human placenta. The probes are useful

XX for antenatal diagnosis of human genetic disorders.

XX Sequence 486 BP; 113 A; 139 C; 104 G; 130 T; 0 other;

XX Query Match 20.8%; Score 486; DB 22; Length 486;
 XX Best Local Similarity 100.0%; Pred. No. 1.8e-129;
 XX Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 570 AAGTATGAAATTCAGCGCAGCAGTGCACACCTCCCGCGTGTGCTCATGTATGTAG 629
 DB 486 AAGTATGAAATTCAGCGCAGCAGTGCACACCTCCCGCGTGTGCTCATGTATGTAG 427
 QY 630 AGTATCAAG 689
 DB 426 AGTATCAAG 367
 QY 690 ATTATCCAGACCGAGCGCGAGTACAGCGCGATCCAGCTCAGCTGAACTGGCAGCGGA 749
 DB 366 ATTATCCAGACCGAGCGCGAGTACAGCGCGATCCAGCTCAGCTGAACTGGCAGCGGA 307
 QY 750 CGAGAGCGCATTCGAAATCATCTACCGGAGAGATCTTTTACTGTGAGGTGCTGCTC 809
 DB 306 CGAGAGCGCATTCGAAATCATCTACCGGAGAGATCTTTTACTGTGAGGTGCTGCTC 247
 QY 810 ACGACTCTTTCGAGGTGCGACCGCGAGATCGGGGTGGCAGAAATCCCACTTCTGCT 869
 DB 246 ACGACTCTTTCGAGGTGCGACCGCGAGATCGGGGTGGCAGAAATCCCACTTCTGCT 187
 QY 870 CAGAGAGAAATCAAGCGCGTGTCCCTTGTCTAATGCTGCACACCACTTACTGTCATG 929
 DB 186 CAGAGAGAAATCAAGCGCGTGTCCCTTGTCTAATGCTGCACACCACTTACTGTCATG 127
 QY 930 GCACCCGGGAATGACTTGGGCAATCACTGACTTGTGTGATTCGCAAGAGACATTTGG 989
 DB 126 GCACCCGGGAATGACTTGGGCAATCACTGACTTGTGTGATTCGCAAGAGACATTTGG 67
 QY 990 GACTGTCTTGAGAAACAGATATGATGATGTTTGTACTTGTCTTCTTCTGAGTTTC 1049
 DB 66 GACTGTCTTGAGAAACAGATATGATGATGTTTGTACTTGTCTTCTTCTGAGTTTC 7
 QY 1050 TGTCTG 1055
 DB 6 TGTCTG 1

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2002, 15:06:10 ; Search time 2586.11 Seconds
(without alignments)
14597.874 Million cell updates/sec

```

Title: US-09-782-953-11
Perfect score: 2331
Sequence: 1 tttttttttccccaggaagt

```

Scoring table: IDENTITY NUC

Searched: 16154066 reqs, 80977433

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum VB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1: em_estba:*

2: em_estchun:*

3: em_estcin:*

4: em_estcmu:*

5: em_estcov:*

6: em_estcpl:*

7: em_estcro:*

8: em_estc:*

9: em_estc1:*

10: gb_est2:*

11: gb_hic:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_eston:*

17: gb_gsa:*

18: em_gsa_hun:*

19: em_gsa_inv:*

20: em_gsa_pln:*

21: em_gsa_vrt:*

22: em_gsa_fun:*

23: em_gsa_man:*

24: em_gsa_mus:*

25: em_gsa_out:*

26: em_gsa_pro:*

27: em_gsa_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query #			DB	ID	Description
	Score	Match	Length			
C 1	939	40.3	972	9	AL571905	AL571905 AL571905
C 2	938.6	40.3	939	9	AL546617	AL546617 AL546617
C 3	919	39.4	959	9	AL567504	AL567504 AL567504
C 4	915.4	39.3	931	9	AL543576	AL543576 AL543576
C 5	908.4	39.0	939	9	AL554886	AL554886 AL554886
C 6	900.6	38.6	938	9	AL544733	AL544733 AL544733

7	896.8	38.5	914	9	AL549888	AL549888
8	889.6	37.3	1041	13	BM450020	BM450020 AGENCOURT
9	864	37.1	906	9	AL544313	AL544313 AGENCOURT
10	855.2	36.7	1014	9	AL559594	AL559594 AGENCOURT
C 11	841.8	36.1	934	9	AL543715	AL543715 AGENCOURT
C 12	828	35.5	909	9	AL577330	AL577330 AGENCOURT
C 13	825	35.4	885	9	AL544755	AL544755 AGENCOURT
C 14	808.6	34.7	849	9	AL566338	AL566338 AGENCOURT
C 15	806.6	34.6	856	12	BG695987	BG695987 AGENCOURT
16	798.2	34.2	837	9	AL124628	AL124628 AGENCOURT
17	793.2	34.0	875	12	BG574693	BG574693 AGENCOURT
18	784.6	33.7	954	14	BQ953935	BQ953935 AGENCOURT
19	779.6	33.4	1058	14	BQ072755	BQ072755 AGENCOURT
20	773	33.4	790	13	BM013950	BM013950 AGENCOURT
21	744.6	31.9	832	9	AL121754	AL121754 AGENCOURT
22	744	31.9	755	13	BI860682	BI860682 AGENCOURT
C 23	741.2	31.8	877	9	AL570491	AL570491 AGENCOURT
C 24	735.8	31.6	750	9	AL575158	AL575158 AGENCOURT
C 25	735.8	31.6	767	9	AL581186	AL581186 AGENCOURT
C 26	733.4	31.5	748	14	BQ015151	BQ015151 UI-H-ED1-
27	728.2	31.2	1078	13	BM541636	BM541636 AGENCOURT
28	725.4	31.1	727	13	BI258232	BI258232 AGENCOURT
29	723.8	31.1	946	14	BQ278576	BQ278576 AGENCOURT
30	710.4	30.5	740	13	BI463566	BI463566 AGENCOURT
31	708.4	30.4	769	13	BI767955	BI767955 AGENCOURT
32	707.8	30.4	718	9	AL131040	AL131040 AGENCOURT
33	707.4	30.3	811	12	BG570239	BG570239 AGENCOURT
34	698	29.9	980	12	BG391816	BG391816 AGENCOURT
35	687.4	29.5	844	12	BG765353	BG765353 AGENCOURT
36	687.2	29.5	800	12	BF437551	BF437551 AGENCOURT
37	676.8	29.0	751	9	AL576189	AL576189 AGENCOURT
38	675	29.0	759	9	AL136125	AL136125 AGENCOURT
39	666.8	28.6	800	12	BG031707	BG031707 AGENCOURT
40	658.4	28.2	751	13	BG328626	BG328626 AGENCOURT
C 41	654.4	28.1	688	14	BM991261	BM991261 UI-H-D10-
C 42	649.8	27.9	692	14	BQ181589	BQ181589 UI-H-E00-
C 43	643.4	27.6	662	13	BI464521	BI464521 AGENCOURT
C 44	642.8	27.6	680	14	BQ018613	BQ018613 AGENCOURT
C 45	640.6	27.5	874	13	BI856992	BI856992 AGENCOURT

ALIGNMENTS

```

RESULT 1
AL5711905/c
LOCUS
DEFINITION
AL5711905 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI029YJ21 3
prime, mRNA sequence.
ACCESSION
AL5711905
VERSION
AL5711905.1 GI:12929663
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 972)
Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
Full-length cDNA libraries and normalization
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1..972
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DI029YJ21"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end

```

enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : filang@lifetech.com URL : <http://fulllength.invitrogen.com>

BASE COUNT 331 a 206 c 161 g 266 t 8 others

ORIGIN

Query Match 40.3%; Score 939; DB 9; Length 972;

Best Local Similarity 98.7%; Pred. No. 5,9e-237; Mismatches 4; Indels 2; Gaps 2;

Matches 959; Conservative

1305 TTGTTTACATCTTTTGGAGAGCAGAAAGCGTTTGGAAAAGTATCACTTTCCCGA 1364
1305 TTGTTTACATCTTTTGGAGAGCAGAAAGCGTTTGGAAAAGTATCACTTTCCCGA 1364
972 TTGTTTACATCTTTTGGAGAGCAGAAAGCGTTTGGAAAAGTATCACTTTCCCGA 913
1365 TTCTCGGATTTCTGACTCTTTTGCACAGACACTTGGGGAACCTTCTCGGAATGGAT 1424
912 TTCTCGGATTTCTGACTCTTTTGCACAGACACTTGGGGAACCTTCTCGGAATGGAT 853
1425 TCACCTAGGATCCCAACCGTGCAGCTGTAACTTGTGCTTTGCAAAAAGAAATTGATCT 1484
852 TCACCTAGGATCCCAACCGTGCAGCTGTAACTTGTGCTTTGCAAAAAGAAATTGATCT 793
1485 GAAATTCCTCTGTAGAAATTTAGCTTATACATTCAGAG-AATAGCATTTCACTGGCAAC 1543
792 GAAATTCCTCTGTAGAAATTTAGCTTATACATTCAGAGAAATAGCAATTCACCTGGCAAC 733
1544 TTTTACGGGAGGAAATTTTATAGGTTTGGGATCGACCTCAGTTTCTGTTT 1603
732 TTTTACGGGAGGAAATTTTATAGGTTTGGGATCGACCTCAGTTTCTGTTT 673
1604 TCTTTTATGTGGTGTCTTATACATGAAATCATAGCCAAAACCTTTTGGAAAAGTGTG 1663
672 TCTTTTATGTGGTGTCTTATACATGAAATCATAGCCAAAACCTTTTGGAAAAGTGTG 613
1664 GTTGAGATAGTGTCTTTTACCCCGAAGACATCAAGATACACTGTAAATTAAGCT 1723
612 GTTGAGATAGTGTCTTTTACCCCGAAGACATCAAGATACACTGTAAATTAAGCT 553
1724 GATGACATATATCTATACCTGTGTACACTGGGTGAAAAGTATGCGAGTGGGACTAA 1783
552 GATGACATATATCTATACCTGTGTACACTGGGTGAAAAGTATGCGAGTGGGACTAA 493
1784 GATGATTTAACTTACCTGTGAAATCATATGTTTGAAGAAAAGCTGTTCCCATGCTTAAAG 1843
492 GATGATTTAACTTACCTGTGAAATCATATGTTTGAAGAAAAGCTGTTCCCATGCTTAAAG 433
1844 GACTTGAATTTAAAGCAATGTAAGTATAGTATGCTGGCATGTGAGAGGATGCA 1903
432 GACTTGAATTTAAAGCAATGTAAGTATAGTATGCTGGCATGTGAGAGGATGCA 373
1904 GTCCTTTCCCATTCATCTCTGATGAAATTTGTTATATACAGTTTAACTTTTGAATTTT 1963
372 GTCCTTTCCCATTCATCTCTGATGAAATTTGTTATATACAGTTTAACTTTTGAATTTT 313
1964 TTTAGTTGAATGTGATGCTGTGTAATAGTATATATTTTGGCCCTTACATTAACCGT 2023
312 TTTAGTTGAATGTGATGCTGTGTAATAGTATATATTTTGGCCCTTACATTAACCGT 253
2024 AACATGTTTGTGATTTGAAATATCTTATGCAAGTAAACATGATGCTTTTGGAAATTT 2083
252 AACATGTTTGTGATTTGAAATATCTTATGCAAGTAAACATGATGCTTTTGGAAATTT 193
2084 GGAAGATGTTTATTTCTTTGAGAGCAAAATATGTTTGCATTAATG-CTTTGATTTTC 2142
192 GGAAGATGTTTATTTCTTTGAGAGCAAAATATGTTTGCATTAATGCTTTGATTTTC 133
2143 ATATCAAGAAATTTGATGAAGCTTCTCAAAAGCTTTTACGGTACTTGTGAAGAGGAGC 2202

Db 132 ATATCAAGAAATTTGATGAAGCTTCTCAAAAGCTTTTACGGTACTTGTGAAGAGGAGC 73
QY 2203 CGGTTTGGGAGAGACCATTCATCGCTGTCAAGTGTCTTTGTTAAGTCTTTTAACT 2262
Db 72 CGGTTTGGGAGAGACCATTCATCGCTGTCAAGTGTCTTTGTTAAGTCTTTTAACT 13
QY 2263 GGAGAGGCTTAACT 2274
Db 12 GGAGAGGCTTAACT 1

RESULT 2
AL546617 939 bp mRNA linear EST 16-FEB-2001

LOCUS AL546617 LTI.NFL006.PL2 Homo sapiens cDNA clone CS001029YJ21 5

DEFINITION prime, mRNA sequence.

ACCESSION AL546617 GI:12879906

VERSION AL546617.1

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 939)

AUTHORS Li, W. B., Gruber, C., Jesse, J. and Polayes, D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished (2001)

COMMENT Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES

source

1. 939

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="CS01029YJ21"

/clone_1bp="LTI.NFL006.PL2"

/issue_type="Placenta"

/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA

was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-stranded cDNA was digested with Not I and

cloned into the Not I and Eco RV sites of the pCMVSPORT 6

vector library was normalized. Library was constructed by

Life Technologies. Contact : Feng Liang Life Technologies,

a division of Invitrogen 9800 Medical Center Drive

Rockville, Maryland 20850, USA Fax : (1) 301 610 8371

Email : filang@lifetech.com URL : <http://fulllength.invitrogen.com>

BASE COUNT 263 a 219 c 238 g 218 t 1 others

ORIGIN

Query Match 40.3%; Score 938.6; DB 9; Length 939;
Best Local Similarity 99.9%; Pred. No. 7.4e-237; Indels 0; Gaps 0;
Matches 938; Conservative

92 ACTAGGGGCTTACCTGCGGGGTCTGTAGCGCTTCACTGTAAAGCAAGATGATTT 151
1 ACTAGGGGCTTACCTGCGGGGTCTGTAGCGCTTCACTGTAAAGCAAGATGATTT 60
152 TAGAACTTTAACTACAGTTTAACTCCCTGATTTGCTGCTGTGCAAAAGTGAATCTT 211
61 TAGAACTTTAACTACAGTTTAACTCCCTGATTTGCTGCTGTGCAAAAGTGAATCTT 120
212 CAGCGAAAGTGAACCAAGGCCCAATTTGAGTCCCTCTTAGAGCTATGACAGGACAT 271
121 CAGCGAAAGTGAACCAAGGCCCAATTTGAGTCCCTCTTAGAGCTATGACAGGACAT 180
272 CACCTTCAGTATTTTAAAGCTTCAAGAGTCAAGATTAACCTTCAAGCAACCTTCTC 331
181 CACCTTCAGTATTTTAAAGCTTCAAGAGTCAAGATTAACCTTCAAGCAACCTTCTC 240
332 CCGAGAGATCCAGGCTCAGTGCATTAAGACTGAGTTTCTGGAAAGAAATGAAGTT 391

Db 241 CGCAGCAGATGCCAGGCTCCAGCTGCATTAAGACTGAGTTCTTGGGAAAGGAAATGAAGTT 300
 QY 392 ATATTTTGTCTCAGACTTACATAGGAAGCTCACACCTGGCTCCGCCAAATCCAGACAA 451
 Db 301 ATATTTTGTCTCAGACTTACATAGGAAGCTCACACCTGGCTCCGCCAAATCCAGACAA 360
 QY 452 GCAGTTTCTGATCTCCCTCCCGCTCCCGCAGTGGATGGAAACAAAGTGGAGATGC 511
 Db 361 GCAGTTTCTGATCTCCCTCCCGCTCCCGCAGTGGATGGAAACAAAGTGGAGATGC 420
 QY 512 GACCCAGTCATAAATCTATGATCTCTTATATGCCATCTCCAAGCTGGCGCCAGGGGAAA 571
 Db 421 GACCCAGTCATAAATCTATGATCTCTTATATGCCATCTCCAAGCTGGCGCCAGGGGAAA 480
 QY 572 GTATGAATTCACGCGAGCGATGACACCACTCCCGAGCTGGTGTGCTCATGTATGTGAGAG 631
 Db 481 GTATGAATTCACGCGAGCGATGACACCACTCCCGAGCTGGTGTGCTCATGTATGTGAGAG 540
 QY 632 TGATCAAG 691
 Db 541 TGATCAAG 600
 QY 692 TATCCAGACCGAGGAGGCGGAGTACACCGCGATCCACCTCAGCTGAACTGGCAGCGACG 751
 Db 601 TATCCAGACCGAGGAGGCGGAGTACACCGCGATCCACCTCAGCTGAACTGGCAGCGACG 660
 QY 752 AGGAGCATTCCTCAATCATCTACCGGAGAGAAATCTTTTACTGTGGAGTGGCTGTGCAC 811
 Db 661 AGGAGCATTCCTCAATCATCTACCGGAGAGAAATCTTTTACTGTGGAGTGGCTGTGCAC 720
 QY 812 GACTTCTTCGGAGTGGCAGCGGAGATCGGGGTGGCAGAAATCCCACTTCACTGTGCTCA 871
 Db 721 GACTTCTTCGGAGTGGCAGCGGAGATCGGGGTGGCAGAAATCCCACTTCACTGTGCTCA 780
 QY 872 GAAGAGATCAAGCGCTGTCCCTTGTCTAATGCTGCACACCACTTCACTGTGCTCA 931
 Db 781 GAAGAGATCAAGCGCTGTCCCTTGTCTAATGCTGCACACCACTTCACTGTGCTCA 840
 QY 932 ACCCGGAATGACTTGGCCCAATCAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 991
 Db 841 ACCCGGAATGACTTGGCCCAATCAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 900
 QY 992 CTGTCTTGAGAAACAGATATGATGTGTTTGTACTT 1030
 Db 901 CTGTCTTGAGAAACAGATATGATGTGTTTGTACTT 939

RESULT 3
 AL567504/c
 LOCUS
 DEFINITION AL567504 LTI_FL013 FBrn1 Homo sapiens cDNA clone CS0DF037YL22 3
 prime, mRNA sequence.
 ACCESSION AL567504
 VERSION AL567504.1
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 AUTHORS Li, W. B., Gruber, C., Jessee, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 Bp 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
 FEATURES
 Location/Qualifiers
 1..959
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS0DF037YL22"
 /clone_lib="LTI_FL013_FBrn1"

/dev_stages="pooled tissue from post conception fetuses (20
 week, 24 week and 26 week)"
 /lab_host="DH10B"
 /note="Organ: Fetal brain; Vector: pCMVSPORT 6; 1st strand
 cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-stranded cDNA was digested with Not I
 and cloned into the Not I and Eco RV sites of the
 pCMVSPORT 6 vector. Library was constructed by Life
 Technologies. Contact : Feng Liang Life Technologies, a
 division of Invitrogen 9800 Medical Center Drive Rockville
 Maryland 20850, USA Fax : (1) 301 610 8371 Email :
 fliang@lifetech.com URL :
 http://fulllength.invitrogen.com"

BASE COUNT 329 a 193 c 162 g 263 t 12 others
 ORIGIN

Query Match 39.4%; Score 919; DB 9; Length 959;
 Best Local Similarity 97.8%; Pred. No. 1.1e-231;
 Matches 938; Conservative 12; Mismatches 7; Indels 2; Gaps 2;

QY 1356 TTTCCAGATTCTCGATTCTGCACTTTGCAACAGCACTTGTTCGGAACTCTTCCT 1415
 Db 959 TTTCCAGATTCTCGATTCTGCACTTTGCAACAGCACTTGTTCGGAACTCTTCCT 900
 QY 1416 GGAATGCATTCACTCAGCAT-CCCCAACCGTGCAACGTTAACTTGTCTTTGCAAAAG 1474
 Db 899 GGAATGCATTCACTCAGCATCCCCAACCGTGCAACGTTAACTTGTCTTTGCAAAAG 840
 QY 1475 AAGTTGATCTGAAATTCCTCTGTAGAATTTAGCTTATACAATTCAGAGNATAGCAGTTTC 1534
 Db 839 AAGTTGATCTGAAATTCCTCTGTAGAATTTAGCTTATACAATTCAGAGNATAGSAG-TTC 781
 QY 1535 ACTGCCAACTTTTAGTGGGTGAGAAATTTAGTTAGTGTTCGGATTCGGACCTCAGTT 1594
 Db 780 ACTGCCAACTTTAGTGGGTGAGAAATTTTAGTTAGTGTTCGGATTCGGACCTCAGTT 721
 QY 1595 TCTGTGTCTTTTATGTGGTGGTTCCTATACATGAATCATAGCCAAAACATTTTGG 1654
 Db 720 TCTGTGTCTTTTATGTGGTGGTTCCTATACATGAATCATAGCCAAAACATTTTGG 661
 QY 1655 AAACGTGTGTGAGTAGTTGGTTCCTTTACCCAGAGACATCAAGATACACATTGTA 1714
 Db 660 AAACGTGTGTGAGTAGTTGGTTCCTTTACCCAGAGACATCAAGATACACATTGTA 601
 QY 1715 AATAAGCTGATAGCATATATTCATACCTCTGTGTACATGATGAGGAAATATGGCAGTG 1774
 Db 600 AATAAGCTGATGWCATATATTCATACCTCTGTGTACATGATGAGGAAATATGGCAGTG 541
 QY 1775 GGAGACTAAGATGATTAACCTACCTGTGAATCATATGTTGTAGGAAAGCTGTCCCAT 1834
 Db 540 GGAGACTAAGATGATTAACCTACCTGTGAATCATATGTTGTAGGAAAGCTGTCCCAT 481
 QY 1835 GTCTAACAGACATGAAATTCAAAGCATGTCAAGTGAATAGATCTGTGGCGATATGAG 1894
 Db 480 GTCTAACAGACATGAAATTCAAAGCATGTCAAGTGAATAGATCTGTGGCGSATATGAG 421
 QY 1895 AGGATGCACTGGCTTTCCCATTCATCTCTGATGAAATGTTTATAGTAAAGCTTAAACATTT 1954
 Db 420 AGGATGCACTGGCTTTCCCATTCATCTCTGATGAAATGTTTATAGTAAAGCTTAAACATTT 361
 QY 1955 GTAATTTTCTTAGTGTGAATGTGTATGTTGGTAAATAGGTATTTATTTGGCCTTA 2014
 Db 360 GTAATTTTCTTAGTGTGAATGTGTATGTTGGTAAATAGGTATTTATTTGGCCTTA 301
 QY 2015 CAATACCGTAACAATGTTGTCAATTTTGAATACATTAATGCCAAGTAACATGCTT 2074
 Db 300 CAATACCGTAACAATGTTGTCAATTTTGAATACATTAATGCCAAGTAACATGCTT 241
 QY 2075 TGGAAATTTGGAAGTGGTTTATTTCTTTGAGAAGCAATATGTTTGCATTAATGCTTT 2134
 Db 240 TGGAAATTTGGAAGTGGTTTATTTCTTTTGAAGAAGCAATATGTTTGCATTAATGCTTT 181
 QY 2135 GATTGTTTCATATCAAGAAATTCATGTAACGTTCTCAAAACCTGTTTACGGTACTTGGTAA 2194

Db 180 GATTGTTTCATATCAAGAAATTGATGAAAGTTTCAAAACCCGTTTACGTAACCTTTGTA 121

Qy 2195 GAGGAGAGCGCGTTTGGAGAGACCAATGTCATGCTGTCGAAGTGTCTTTGTTAAGTCT 2254

Db 120 GAGGAGAGCGCGTTTGGAGAGACCAATGTCATGCTGTCGAAGTGTCTTTGTTAAGTCT 61

Qy 2255 TTTAACTGAGAGGCTAACCTCAAAATATCTTTTAACTGATCTTATTAATAATG 2313

Db 60 TTTAACTGAGAGGCTAACCTCAAAATATCTTTTAACTGATCTTATTAATAATG 2

RESULT 4
AL543576 931 bp mRNA linear EST 16-FEB-2001
LOCUS AL543576 LTI_NFL006.PL2 Homo sapiens cDNA clone CS0D106YK09 5
DEFINITION prime, mRNA sequence.
ACCESSION AL543576 GI:12876055
VERSION AL543576
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 931)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
JOURNAL
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web: www.genoscope.cns.fr.
Location/Qualifiers
1. 931
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D106YK09"
/clone_1lb="LTI_NFL006.PL2"
/ribose_type="placenta"
/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact: Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive, Rockville, Maryland 20850, USA Fax: (1) 301 610 8371
Email: fliang@life.com URL: http://fulllength.invitrogen.com"

BASE COUNT 262 a 217 c 238 g 214 t

ORIGIN

Query Match 39.3%; Score 915.4; DB 9; Length 931;
Best Local similarity 99.8%; Pred. No. 1e-230; Indels 1; Gaps 1;
Matches 927; Conservative 0; Mismatches 1;

Qy 95 AGGGCTTGACTGGGTGCTGTAGCCCTTCACTGTAGAAAGCAAGATGCAATTTAG 154

Db 1 AGGGCTTGACTGGGTGCTGTAGCCCTTCACTGTAGAAAGCAAGATGCAATTTAG 60

Qy 155 AAACCTTAACTACAGTTTAACTCCCTGATTCCTGTGGCAACAGTATATCTTACG 214

Db 61 AAACCTTAACTACAGTTTAACTCCCTGATTCCTGTGGCAACAGTATATCTTACG 120

Qy 215 CGAAAGTGAACACGAGGCCAAATTTGAGTCCCTTTAGAGAGTATGACAGACATAC 274

Db 121 CGAAAGTGAACACGAGGCCAAATTTGAGTCCCTTTAGAGAGTATGACAGACATAC 180

Qy 275 CTTTCAGTATTTTAAAGAGCTTCAACGAGTCAAGATTAATCTTACAGCAACCCCTTCCGC 334

Db 181 CTTTCAGTATTTTAAAGAGCTTCAACGAGTCAAGATTAATCTTACAGCAACCCCTTCCGC 240

Qy 335 AGCAGATGCCAGGCTTCAGCTGATAGACTGATTTCTGGGAAAGAAATGAATTA 394

Db 241 AGCAGATGCCAGGCTTCAGCTGATAGACTGATTTCTGGGAAAGAAATGAATTA 300

Qy 395 TTTTGTCTGAGACCTTACATATGAAAGCTCACACCTGCTCCGCCAAATCCAGACAAGA 454

Db 301 TTTTGTCTGAGACCTTACATATGAAAGCTCACACCTGCTCCGCCAAATCCAGACAAGA 360

Qy 455 GTTTGTGATCTCCCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCG 514

Db 361 GTTTGTGATCTCCCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCG 420

Qy 515 CCCAGTATTAATATGATCTTATATGCAATCTCCAGAGCTGGGGCCAGGGGAAAGTA 574

Db 421 CCCAGTATTAATATGATCTTATATGCAATCTCCAGAGCTGGGGCCAGGGGAAAGTA 480

Qy 575 TGAATTGACAGAGCCGACTGACACCACTCCAGCGTGTGTGTCATGATATGAGAGTA 634

Db 481 TGAATTGACAGAGCCGACTGACACCACTCCAGCGTGTGTGTCATGATATGAGAGTA 540

Qy 635 TCAGAGAGAGAGAGAGAGAGAGAAATGGAATGAGAGACCTTAAGCCAAATTTAT 694

Db 541 TCAGAGAGAGAGAGAGAGAGAGAAATGGAATGAGAGACCTTAAGCCAAATTTAT 600

Qy 695 CCAGACCAAGAGCGCGAGATACAGCCGATCCACTGAGCTGAATGACGACGAGAG 754

Db 601 CCAGACCAAGAGCGCGAGATACAGCCGATCCACTGAGCTGAATGACGACGAGAG 660

Qy 755 AGCATTCCAATCATATCTACCGGAGAAATCTTTACTGTGAGAGTGTGCTGACGAC 814

Db 661 AGCATTCCAATCATATCTACCGGAGAAATCTTTACTGTGAGAGTGTGCTGACGAC 720

Qy 815 TTTCTGAGAGTGGCAGCCGAGATCCGGGGTGGAGAAATCCAGTTATGTTGCTCGAA 874

Db 721 TTTCTGAGAGTGGCAGCCGAGATCCGGGGTGGAGAAATCCAGTTATGTTGCTCGAA 780

Qy 875 GAGATCAAGGCGGTCCTGCTTCTTAATGCTGACACACAGTACTGTTCAATGGCACC 934

Db 781 GAGATCAAGGCGGTCCTGCTTCTTAATGCTGACACACAGTACTGTTCAATGGCACC 840

Qy 935 CGGGAATGACTTGGGCAATCATGAGTTGTGTGATCGCAAGACATTTGGGACTG 994

Db 841 CGGGAATGACTTGGGCAATCATGAGTTGTGTGATCGCAAGACATTTGGGACTG 900

Qy 995 TTTTGAAGAAACAGTATATGATGCTT 1023

Db 901 TCTTGAG-AAACAGATATGATGATTT 928

RESULT 5
AL554686 939 bp mRNA linear EST 16-FEB-2001
LOCUS AL554686 LTI_NFL006.PL2 Homo sapiens cDNA clone CS0D1085YF09 5
DEFINITION prime, mRNA sequence.
ACCESSION AL554686
VERSION AL554686 GI:12895707
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 939)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
JOURNAL
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web: www.genoscope.cns.fr.
Location/Qualifiers
1. 939
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1085YF09"

```

/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/notes="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT      262 a  219 c  238 g  218 t      2 others
ORIGIN

```

```

Query Match      39.0%; Score 908.4; DB 9; Length 939;
Best Local Similarity 99.3%; Pred. No. 7.1e-229;
Matches 933; Conservative 1; Mismatches 3; Indels 3; Gaps 2;

QY  92  ACTAGGGGCTTCACTCGCTGGCTCTGTAGCGCTTTCACGTGAAGAAACCAAGATGCAATTT 151
DB  1  ACTAGGGGCTTCACTCGCTGGCTCTGTAGCGCTTTCACGTGAAGAAACCAAGATGCAATTT 60

QY 152  TAGAAACTTTAACTACAGTTTTAGCTCCCTCGATTCCTGCTGCTGGCAACACGTATATCTT 211
DB  61  TAGAAACTTTAACTACAGTTTTAGCTCCCTCGATTCCTGCTGCTGGCAACACGTATATCTT 120

QY 212  CAGCGAAAGTGAACCA--GGCCMAATTTGAGTCCCTCTTTAGGACGTATGCAAGGAC 269
DB  121  CAGCGAAAGTGAACCAATGTCGCCAAATTTGAGTCCCTCTTTAGGACGTATGCAAGGAC 180

QY 270  ATCACTTTTCAGTATTTTAAAGAGCTTCAACAGCTCAGATTAACATTCAGAACCCCTTC 329
DB  181  ATCACTTTTCAGTATTTTAAAGAGCTTCAACAGCTCAGATTAACATTCAGAACCCCTTC 240

QY 330  TCCGACAGAGATGCGAGCTCCAGCTGCATAGAAGCTGAGTTTCTGGAAAGGAAATGAAG 389
DB  241  TCCGACAGAGATGCGAGCTCCAGCTGCATAGAAGCTGAGTTTCTGGAAAGGAAATGAAG 300

QY 390  TTATATTGCTCAGACCTTACATAGGAAGCTCACACCTGGCTCGCCAAATCCAGAC 449
DB  301  TTATATTGCTCAGACCTTACATAGGAAGCTCACACCTGGCTCGCCAAATCCAGAC 360

QY 450  AAGCAGTTTCGATTCCTCCCTCCGCTCTCCGCTAGTGGATGGAACAAAGTGAAGAT 509
DB  361  AAGCAGTTTCGATTCCTCCCTCCGCTCTCCGCTAGTGGATGGAACAAAGTGAAGAT 420

QY 510  GCGACCCAGTCATAACTATGATCTCTTATATGCCATCTCCAAGCTGGGGCCAGGGAA 569
DB  421  GCGACCCAGTCATAACTATGATCTCTTATATGCCATCTCCAAGCTGGGGCCAGGGAA 480

QY 570  AAGTATGAATTCGACGAGGAGCTGACACCACTCCGAGCTGGTGTCCATGATGTGAG 629
DB  481  AAGTATGAATTCGACGAGGAGCTGACACCACTCCGAGCTGGTGTCCATGATGTGAG 540

QY 630  AGTGATCAAGAGAGGAGGAGGAGGAAATGGAAGAAATGAGGAGACTTAAGCCAAA 689
DB  541  AGTGATCAAGAGAGGAGGAGGAGGAAATGGAAGAAATGAGGAGACTTAAGCCAAA 600

QY 690  ATTATCCAGACCAAGGAGGCGGAGTACACCGGATCCACTCAGCTGAACTGCCACGGA 749
DB  601  ATTATCCAGACCAAGGAGGCGGAGTACACCGGATCCACTCAGCTGAACTGCCACGGA 660

QY 750  CGAGGACGCAATTCAAATCATACTCAGCGGAGGAATCTTTTCTGAGGCTGGCTGTC 809
DB  661  CGAGGACGCAATTCAAATCATACTCAGCGGAGGAATCTTTTCTGAGGCTGGCTGTC 720

QY 810  ACCACTTCTTCGAGTGGCAGCGAGATCGGGGTGGCAGAAATCCCAATTCATGCTCT 869
DB  721  ACCACTTCTTCGAGTGGCAGCGAGATCGGGGTGGCAGAAATCCCAATTCATGCTCT 780

QY 870  CAGAAGAGAAATCAAGGCGGTGTCCTTCTTAATGCTGCACACCACTGCTGTCATG 929

```

```

DB 781  CAGAGAGAAATCAAGCGCGTGTCCCTTGTCTTAATGCTGCACACCACTTACTGTCATG 840
QY 930  GCACCCGGGAATGACTTGGGCCAATCACTGAGTTTGTGTGATCGCACAGGACATTTGG 989
DB 841  GCACCCGGGAATGACTTGGGCCAATCACTGAGTTTGTGTGATCGCACAGGACATTTGG 900
QY 990  GACTGCTCTGAGAAACAGATAATGATAGTGTGTGTACT 1029
DB 901  GACTGCTCTGAG-AAAACAGATAATGATAGTGTGTGTACT 939

RESULT 6
AL544733/c
LOCUS      AL544733      938 bp      mRNA      linear      EST 16-FEB-2001
DEFINITION AL544733 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI022VJ23 3
prime, mRNA sequence.
ACCESSION  AL544733
VERSION    AL544733.1 GI:12877213
KEYWORDS   EST.
SOURCE      human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 938)
AUTHORS   Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE     Full-length cDNA libraries and normalization
JOURNAL   Unpublished (2001)
COMMENT   Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES   Location/Qualifiers
            1..938
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone_lib="LTI_NFL006_PL2"
            /tissue_type="placenta"
            /notes="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA
            was primed with a NotI-oligo(dT) primer. Five prime end
            enriched, double-stranded cDNA was digested with Not I and
            cloned into the Not I and Eco RV sites of the pCMVSPORT 6
            vector. Library was normalized. Library was constructed by
            Life Technologies. Contact : Feng Liang Life Technologies,
            a division of Invitrogen 9800 Medical Center Drive
            Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
            Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com"

BASE COUNT      325 a  194 c  153 g  263 t      3 others
ORIGIN

```

```

Query Match      38.6%; Score 900.6; DB 9; Length 938;
Best Local Similarity 98.9%; Pred. No. 8.2e-227;
Matches 914; Conservative 2; Mismatches 7; Indels 1; Gaps 1;

QY 1371 GATTCCTGACTCTTTGCAACAGCACTTGTTCGGGAATCTTCTCTGGAATGCATTCAC 1430
DB 930  GAATCTCRACTCTTTACACAGAACTTGTTCGGGAATCTTCTCTGGAATGCATTCAC 871

QY 1431 AGCATCCCCAACCGTGCACGTGTAACTTGTGCTTTTGCAAAAGAGTTGATCTGAAAT 1490
DB 870  AGCATCCCCAACCGTGCACGTGTAACTTGTGCTTTTGCAAAAGAGTTGATCTGAAAT 811

QY 1491 CCTCTGTAGAAATTTAGCTTATACAATTCAGAGAAATGAGCTTTCATCGCAACTTTAGT 1550
DB 810  CCTCTGTAGAAATTTAGCTTATACAATTCAGAGAAATGAGCTTTCATCGCAACTTTAGT 752

QY 1551 GGTGTGAAATTTAGTTTGTGATCGGACCTCGACCTCAGTTCTCTGTTGTTCTTTTA 1610
DB 751  GGTGTGAAATTTAGTTTGTGATCGGACCTCGACCTCAGTTCTCTGTTGTTCTTTTA 692

QY 1611 TGTGTGTTTCTTATACATGAATCATAGCCAAATCTTTTGGAAACCTGTTGGTGA 1670

```

Db 691 TGTGTGCTTCTATACATGAATCATACCAAAAACCTTTTGGAACTGTTGGTGA 632
 QY 1671 TAGTGGTCTTTTACCCACGAGACATCAAGATACATCTGTAATTAAGCTGATACA 1730
 Db 631 TAGTGGTCTTTTACCCACGAGACATCAAGATACATCTGTAATTAAGCTGATACA 572
 QY 1731 TATATTCATACCTGTTGTAACATCTGGGTGAAAAGTAAGGAGGAGGAGTAAAGTAT 1790
 Db 571 TATATTCATACCTGTTGTAACATCTGGGTGAAAAGTAAGGAGGAGGAGTAAAGTAT 512
 QY 1791 TAACTCACTGTAATCATATGTTAGGAAAAGCTGTTCCCATGCTTAACAGACTTGA 1850
 Db 511 TAACTCACTGTAATCATATGTTAGGAAAAGCTGTTCCCATGCTTAACAGACTTGA 452
 QY 1851 ATTCAAAGCATGCAAGTGAATAGATCTGTGGCATATAGAGGAGTGAAGTCCCTT 1910
 Db 451 ATTCAAAGCATGCAAGTGAATAGATCTGTGGCATATAGAGGAGTGAAGTCCCTT 392
 QY 1911 TCCCATTCATCTGATGGAATGTTATCTAGTTACATTTGTAATTTTTCTAGT 1970
 Db 391 TCCCATTCATCTGATGGAATGTTATCTAGTTACATTTGTAATTTTTCTAGT 332
 QY 1971 TGTATGTGATGCTGTGAATAGTATTAATTTGGCTTACAAATACCTTAACATG 2030
 Db 331 TGTATGTGATGCTGTGAATAGTATTAATTTGGCTTACAAATACCTTAACATG 272
 QY 2031 TTTGTCTTTTGAATTAATTAATGCTTAATTAATTTGGCTTACAAATACCTTAACATG 2090
 Db 271 TTTGTCTTTTGAATTAATTAATGCTTAATTAATTTGGCTTACAAATACCTTAACATG 212
 QY 2091 GGTTTATCTGTTGAGAAAGCAATATGTTGATTAATGCTTGAATTTGTTATCAAG 2150
 Db 211 GGTTTATCTGTTGAGAAAGCAATATGTTGATTAATGCTTGAATTTGTTATCAAG 152
 QY 2151 AAATGATTTGAACCTTCTCAACCTGTTTACGTAATTTGTAAGAGGAGCCGGTTGG 2210
 Db 151 AAATGATTTGAACCTTCTCAACCTGTTTACGTAATTTGTAAGAGGAGCCGGTTGG 92
 QY 2211 GAGGACCATTTGCAATGCTGTGCAAGGTTTCTGTTAAGTGTCTTTAACTGAGAGGC 2270
 Db 91 GAGGACCATTTGCAATGCTGTGCAAGGTTTCTGTTAAGTGTCTTTAACTGAGAGGC 32
 QY 2271 TAACTCAAAATACCTTTTAACT 2294
 Db 31 TAACTCAAAATATTTTTTTTAC 8

RESULT 7 914 bp mRNA linear EST 16-FEB-2001
 AL549888
 LOCUS AL549888 LTI_NFL006_PU2 Homo sapiens cDNA clone CS0D1054YB17 5
 DEFINITION prime: mRNA sequence.
 ACCESSION AL549888
 VERSION AL549888.1 GI:12886309
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 914)
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 CONTACT Genoscope
 COMMENT Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
 source 1..914
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS0D1054YB17"
 /clone_1ib="LTI_NFL006_PU2"

/cissue_type="placenta"
 /note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact: Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@life.com URL : http://fulllength.invitrogen.com"

BASE COUNT 238 a 166 c 209 g 300 t 1 others
 ORIGIN
 Query Match 38.5%; Score 896.8; DB 9; Length 914;
 Best Local Similarity 99.5%; Pred. No. 8.2e-226;
 Matches 909; Conservative 1; Mismatches 3; Indels 1; Gaps 1;
 QY 956 ACTGAGTTTGTGATTCGCAAGACATTTGGGACTGTTGAGAAAACAGATATGA 1015
 Db 1 ACTGAGTTTGTGATTCGCAAGACATTTGGGACTGTTGAGAAAACAGATATGA 60
 QY 1016 TAGTGTGTTGACTGTTCTTTCTGGTAGTTCTGTGTGCCAAGGCGAGTTGATCA 1075
 Db 61 TAGTGTGTTGACTGTTCTTTCTGGTAGTTCTGTGTGCCAAGGCGAGTTGATCA 120
 QY 1076 GTGAGCTCAGAGAGACCTTCTGTTCTTAAGTGGCTGCAAGGCGCACTCTACTGGT 1135
 Db 121 GTGAGCTCAGAGAGACCTTCTGTTCTTAAGTGGCTGCAAGGCGCACTCTACTGGT 180
 QY 1136 AGGAAGAGGTACCAAGAGAGCCGCTAGTACAGAGAGTGTGAAAACAGACATATG 1195
 Db 181 AGGAAGAGGTACCAAGAGAGCCGCTAGTACAGAGAGTGTGAAAACAGACATATG 240
 QY 1196 AATGTGAATTTGAGCGTTTCTTTCTCCCTCATGTTCTCATGTTTGCATGTAT 1255
 Db 241 AATGTGAATTTGAGCGTTTCTTTCTCCCTCATGTTCTCATGTTTGCATGTAT 300
 QY 1256 TACTGATTTACAGACTTAACCTTTGTTGTAATTAAGTACACGTTGTTTACAT 1315
 Db 301 TACTGATTTACAGACTTAACCTTTGTTGTAATTAAGTACACGTTGTTTACAT 360
 QY 1316 CTTTGGGAGGAGGAGGAGGTTTGGAAAAGTATCACTTTCCAGATTCTCGAATTC 1375
 Db 361 CTTTGGGAGGAGGAGGAGGTTTGGAAAAGTATCACTTTCCAGATTCTCGAATTC 420
 QY 1376 TCGACTCTTTGCAACAGCATTTGCTGCGAACTTCTCTGGAATGCACTTACAGAT 1435
 Db 421 TCGACTCTTTGCAACAGCATTTGCTGCGAACTTCTCTGGAATGCACTTACAGAT 480
 QY 1436 CCCCAACCGTGAAGGTGAATCTGTGCTTTTGCAAAAGAGTTGATGAAATTCCTCT 1495
 Db 481 CCCCAACCGTGAAGGTGAATCTGTGCTTTTGCAAAAGAGTTGATGAAATTCCTCT 540
 QY 1496 GTAGAAATTTAGTTATACATTCAGAGAAATAGCACTTCACTCCCACTTTAGTGG 1555
 Db 541 GTAGAAATTTAGTTATACATTCAGAGAAATAGCACTTCACTCCCACTTTAGTGG 600
 QY 1556 AGAAATTTAGTTAGTGTGTTGGATCGGACCTCAAGTCTGTTGTTCTTTATG 1615
 Db 601 AGAAATTTAGTTAGTGTGTTGGATCGGACCTCAAGTCTGTTGTTCTTTATG 660
 QY 1616 TGGTTCTATACATGATCATAGCCAAAACCTTTTGGAACTGTGGTGAATAGT 1675
 Db 661 TGGTTCTATACATGATCATAGCCAAAACCTTTTGGAACTGTGGTGAATAGT 720
 QY 1676 GGTTCCTTTTACCCACGAAAGCATCAAGTACATTTGTAATTAAGTGAATAT 1735
 Db 721 GGTTCCTTTTACCCACGAAAGCATCAAGTACATTTGTAATTAAGTGAATAT 780
 QY 1736 TCATACCTGTTGACCTTGGGTGAAGATGAGGAGTGGAGACTTAAGATTAAC 1794
 Db 781 TCATACCTGTTGACCTTGGGTGAAGATGAGGAGTGGAGACTTAAGATTAAC 840

```

QY 1795 CTACCTGTGAATCATATGTTGTAGGAAAGCTGTTCCCATGCTTAACAGGACTTGAATTC 1854
Db 841 CTACCTGTGAATCATATGTTGTAGGAAAGCTGTTCCCATGCTTAACAGGACTTGAATTC 900
QY 1855 AAAGCATGTCAACT 1868
Db 901 AAAGCATGTTAAGT 914

RESULT 8
BM450020
LOCUS
DEFINITION
  BM450020 1041 bp mRNA linear EST 05-FEB-2002
  AGENCOURT_6393352 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5528025
  5', mRNA sequence.
ACCESSION
  BM450020
VERSION
  BM450020.1 GI:18499060
KEYWORDS
  EST.
SOURCE
  human.
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 1041)
  NIH-MGC http://mgc.nci.nih.gov/.
  National Institutes of Health, Mammalian Gene Collection (MGC)
  Unpublished (1999)
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-remail.nih.gov
  Tissue Procurement: ATCC/DCTP
  cDNA Library Preparation: Life Technologies, Inc.
  cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
  DNA Sequencing by: Agencourt Bioscience Corporation
  Clone distribution: MGC clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  http://image.llnl.gov
  Plate: LHAM12204 row: d column: 10
  High quality sequence stop: 643.
  Location/Qualifiers
    1..1041
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /clone="IMAGE:5528025"
      /clone_lib="NIH MGC 72"
      /tissue_type="melanotic melanoma"
      /lab_host="DH10B (phage-resistant)"
      /note="Organ: skin; Vector: pCMV-Sport6; Site 1: NotI;
      Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
      Average insert size 2 kb. Library constructed by Life
      Technologies."
  BASE COUNT 282 a 248 c 274 g 236 t 1 others
  ORIGIN

Query Match 37.3%; Score 869.6; DB 13; Length 1041;
Best Local Similarity 98.7%; Pred. No. 1.3e-218;
Matches 898; Conservative 0; Mismatches 9; Indels 3; Gaps 2;

QY 61 AGGAACCTACAGCCTCTTGGAAGGAATCTCACTAGGGCTTGACGCGGTCTGTAG 120
Db 1 AGGAACCTACAGCCTCTTGGAAGGAATCTCACTAGGGCTTGACGCGGTCTGTAG 60
QY 121 CGCTTTCACTGTAAGAACCAAGATGATCTTCAAGCGAAGTGAACACCGGGCCAAATTG 180
Db 61 CGCTTTCACTGTAAGAACCAAGATGATCTTCAAGCGAAGTGAACACCGGGCCAAATTG 120
QY 181 TGAATGCTGTGTGGCAACAGTGAATCTTCAAGCGAAGTGAACACCGGGCCAAATTG 240
Db 121 TGAATGCTGTGTGGCAACAGTGAATCTTCAAGCGAAGTGAACACCGGGCCAAATTG 180
QY 241 AGTCCCTCTTTAGGACGTATGACAGGACATCACCTTTCAAGTATTTAAGACTTCAAC 300
Db 181 AGTCCCTCTTTAGGACGTATGACAGGACATCACCTTTCAAGTATTTAAGACTTCAAC 240
QY 301 GAGTCAGATAAATTTTCAGCAACCCCTTCTCGGAGCAGATGCCAGGCTCCAGCTGCATA 360

```

```

Db 241 GAGTCAGATAAATTTTCAGCAACCCCTTCTCGGAGCAGATGCCAGGCTCCAGCTGCATA 300
QY 361 AGACTGAGTTTCTGGGAAGGAATGAAGTATATTTTGTCTCAGACTTACACATAGGAA 420
Db 301 AGACTGAGTTTCTGGGAAGGAATGAAGTATATTTTGTCTCAGACTTACACATAGGAA 360
QY 421 GCTCACACCTTGCTCCGCCAAATCCAGACAAGCAGTTTCTGATCTCCCTCCCGCTCTC 480
Db 361 GCTCACACCTTGCTCCGCCAAATCCAGACAAGCAGTTTCTGATCTCCCTCCCGCTCTC 420
QY 481 CGCAGTGGGATGGAACAAGTGAAGATGCGACCCAGTCAATAACTATGATCTCTTAT 540
Db 421 CGCAGTGGGATGGAACAAGTGAAGATGCGACCCAGTCAATAACTATGATCTCTTAT 480
QY 541 ATGCCATCTCCAAGCTGGGGCCAGGGGAAAGTATGAATTGCACGCGACTGACACCA 600
Db 481 ATGCCATCTCCAAGCTGGGGCCAGGGGAAAGTATGAATTGCACGCGACTGACACCA 540
QY 601 CTCCAGCGTGGTGCATGTATGTGAGATGATCAAGAGAAGGAGGAAGAGGAAA 660
Db 541 CTCCAGCGTGGTGCATGTATGTGAGATGATCAAGAGAAGGAGGAAGAGGAAA 600
QY 661 TGGAAAGAAATGAGGAGACCTTAAGCCAAAATTTATCCAGACAGGAGGCGGAGTACAGC 720
Db 601 TGGAAAGAAATGAGGAGACCTTAAGCCAAAATTTATCCAGACAGGAGGCGGAGTACAGC 660
QY 721 CGATCACCTCAGCTGAATCGCAGCGACGAGGACGATTCMAATCATATCAACGGGA 780
Db 661 CGATCACCTCAGCTGAATCGCAGCGACGAGGACGATTCMAATCATATCAACGGGA 720
QY 781 GGAATCTTTACTGTTGGAGTGGTGTGTCAGACTTCTTCGGAGTGGCAGCGAGATCG 840
Db 721 GGAATCTTTACTGTTGGAGTGGTGTGTCAGACTTCTTCGGAGTGGCAGCGAGATCG 780
QY 841 GGGTGGCAAAATCCAGTTCATGTTGCTCAGAGAAGATCAAGGCGGTGTCCTTTGTT 900
Db 781 GGGTGGCAAAATCCAGTTCATGTTGCTCAGAGAAGATCAAGGCGGTGTCCTTTGTT 840
QY 901 CTAATGCTGCACA-CCAGTACTGTTCAATGACCCCGG--AATGACTTGGGCCAATCAC 957
Db 841 CTAATGCTGCACACCCAGTTACTGTTCAATGACCCCGGAAATGACTTGGGCCAATCAC 900
QY 958 TCAGTTGTG 967
Db 901 TGGAGTTTG 910

RESULT 9
AL544313
LOCUS
DEFINITION
  AL544313 LTI_NFL006_P12 Homo sapiens cDNA clone CSODI019YH04 5
  Prime, mRNA sequence.
ACCESSION
  AL544313
VERSION
  AL544313.1 GI:12876793
KEYWORDS
  EST.
SOURCE
  human.
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 906)
  Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
  Full-length cDNA libraries and normalization
  Unpublished (2001)
  Contact: Genoscope
  Genoscope - Centre National de Sequencage
  BP 191 91006 EVRY cedex - France
  Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr.
  Location/Qualifiers
    1..906
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /clone="CSODI019YH04"
  FEATURES
    source

```


/clone_lib="LTI_NFL006_PL2"
 /tissue_type="placenta"
 /note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
 Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com"

BASE COUNT 253 a 215 c 229 g 202 t 7 others
 ORIGIN

Query Match 37.1%; Score 864; DB 9; Length 906;
 Best Local Similarity 98.7%; Pred. No. 3.9e-217;
 Matches 894; Conservative 2; Mismatches 2; Indels 3; Gaps 3;

OY 100 CTGACTGCGTGGGTCTGTAGCGCTTCACTGTGAAGAAAGCAAGATGATTTAGAAACT 159
 DB 1 CTGACTGCGTGGGTCTGTAGCGCTTCACTGTGAAGAAAGCAAGATGATTTAGAAACT 60
 OY 160 TTACTACAGATTAGTCTCCCTGATTGCTGTGTGGCAACAGTATATCTTCAGCGAAA 219
 DB 61 TTACTACAGATTAGTCTCCCTGATTGCTGTGTGGCAACAGTATATCTTCAGCGAAA 120
 OY 220 GTGAAACGAGGCGCAATTTGAGTCCCTCTTTAGACATGACAGACATCACTTTC 279
 DB 121 GTGAAACGAGGCGCAATTTGAGTCCCTCTTTAGACATGACAGACATCACTTTC 180
 OY 280 AGATTTTAAAGCTTCAACGAGTCAAGTAATTAATTCAGCAACCCCTCTCCGACAG 339
 DB 181 AGATTTTAAAGCTTCAACGAGTCAAGTAATTAATTCAGCAACCCCTCTCCGACAG 240
 OY 340 ATGCGACGCTCAGCTTCATAGACTGATTTCTGGAAAGAAATGAAGTTATTTTG 399
 DB 241 ATGCGACGCTCAGCTTCATAGACTGATTTCTGGAAAGAAATGAAGTTATTTTG 300
 OY 400 CTCAGACCTTACATAGGAAGCTCACAACCTGCTCCGCCAATTCAGACAGACTTTC 459
 DB 301 CTCAGACCTTACATAGGAAGCTCACAACCTGCTCCGCCAATTCAGACAGACTTTC 360
 OY 460 TGATCTCCCTCCGCTCTCCGCACTGAGTGAAGAAAGTGAAGATGCAACCCAG 519
 DB 361 TGATCTCCCTCCGCTCTCCGCACTGAGTGAAGAAAGTGAAGATGCAACCCAG 420
 OY 520 TCATTAACATGATCTCTTATATGCACTCTCAAGCTGGGGCCAGGGGAAAGATGAAT 579
 DB 421 TCATTAACATGATCTCTTATATGCACTCTCAAGCTGGGGCCAGGGGAAAGATGAAT 480
 OY 580 TGCACGCGACGACGACCACTCCACGCTGTGTGTCATGATGTGAGAGTATCAAG 639
 DB 481 TGCACGCGACGACGACCACTCCACGCTGTGTGTCATGATGTGAGAGTATCAAG 539
 OY 640 AGAAGAGGAAGAAAGGAATGGAAGAAATGAGGAGACTTAAGCAAAATTTTCCAGA 699
 DB 540 AGAAGAGGAAGAAAGGAATGGAAGAAATGAGGAGACTTAAGCAAAATTTTCCAGA 599
 OY 700 CCAGAGGCGGAGATGACACGCCGACCTCAAGCTGAGCAAGCGGAGAGAGAGCA 759
 DB 600 CCAGAGGCGGAGATGACACGCCGACCTCAAGCTGAGCAAGCGGAGAGAGAGCA 659
 OY 760 TTCCAATCATCTACGCGAGGAATCTTTTACTGTGAGGTGGCTGTGCAAGACTTCTT 819
 DB 660 TTCCAATCATCTACGCGAGGAATCTTTTACTGTGAGGTGGCTGTGCAAGACTTCTT 719
 OY 820 CCGAGGTGCGACCGAGATCGGGGTGGCAGAAATCCCACTTATATTTCTCAGAAAGAA 879
 DB 720 CCGAGGTGCGACCGAGATCGGGGTGGCAGAAATCCCACTTATATTTCTCAGAAAGAA 779
 OY 880 TCAAGGCGGTGCTCTTCTTCTAATGTGACACCAAGTATCTGTTCATGCGACCCGGA 939

DB 780 TCAAGGCGGTGCTCTTCTTCTAATGCTACACCACTACTGTTCATGCA-CCGGA 838
 OY 940 ATGACTGGGCGCAATCACTAGTTGTGTGATTCGACAGACATTTGGACTGTCTTG 999
 DB 839 ATGACTGGGCGCAATCACTAGTTGTGTGATTCGACAGACATTTGGACTGTCTTG 897
 OY 1000 AGAAAA 1005
 DB 898 AGAAAA 903

RESULT 10
 AL559594 1014 bp mRNA linear EST 16-FEB-2001
 LOCUS AL559594 LTI_NFL008.TC2 Homo sapiens cDNA clone CS0D013YN24 5
 DEFINITION prime, mRNA sequence.
 ACCESSION AL559594
 VERSION AL559594.1 GI:12905231
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1014)
 Li, W.B., Gruber, C., Jesse, J., and Polayes, D. Full-length cDNA libraries and normalization Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France
 Email: seque@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
 source
 1..1014
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS0D013YN24"
 /clone_lib="LTI_NFL008.TC2"
 /sex="male"
 /tissue_type="T cells from T cell leukemia"
 /note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
 Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com"

BASE COUNT 271 a 237 c 270 g 234 t 2 others
 ORIGIN

Query Match 36.7%; Score 855.2; DB 9; Length 1014;
 Best Local Similarity 96.7%; Pred. No. 8.5e-215;
 Matches 913; Conservative 2; Mismatches 25; Indels 4; Gaps 4;

OY 196 CAAACGATATCTTTCAGCGCAAGTGAACCGAGGCCCAATTTGAGTCCCTTTAGCA 255
 DB 51 CAATCTGCACTATTGGAGAAAGCAATGTTGGCCCAATTTGAGTCCCTTTAGCA 110
 OY 256 CGTATGACAGACATCACTTTCAGTATTTTAAAGCTTCAAGAGATCAAGATTAAC 315
 DB 111 CGTATGACAGACATCACTTTCAGTATTTTAAAGCTTCAAGAGATCAAGATTAAC 170
 OY 316 TCGCAACCCCTTCTCCGAGAGATGCCAGGCTCAGCTGCATTAAGACTGAGTTCTGG 375
 DB 171 TCGCAACCCCTTCTCCGAGAGATGCCAGGCTCAGCTGCATTAAGACTGAGTTCTGG 230
 OY 376 GAAAGGAATGAAGTATATTTTGTCTCAGACCTTACATAGGAAGTCAACCTGGCTC 435
 DB 231 GAAAGGAATGAAGTATATTTTGTCTCAGACCTTACATAGGAAGTCAACCTGGCTC 290
 OY 436 CGCAATTCAGACAGAGTTTGTATCTCCCTCCGCTCTCCGCAAGTGGATGGA 495

Db 128 GATTGTCATATCAAGAAATGATGAACGTTCTCAAAACCTGTTTACGGTACTGTGTA 69
 QY 2195 GAGGAGCCGCTTTGGGAGAACCATGTCAGTGCCTCAAGTGTCTTCTTTAGTACT 2254
 Db 68 GAGGAGCCGCTTTGGGAGAACCATGTCAGTGCCTGCHCAAGTGTCTTCTTTAGTACT 9
 QY 2255 TTTAACT 2262
 Db 8 TTTAACT 1

RESULT 12
 AL577330 909 bp mRNA linear EST 16-FEB-2001
 LOCUS AL577330 LTI_NFL006.PL2 Homo sapiens cDNA clone CS0D1085F09.3
 DEFINITION prime, mRNA sequence.
 ACCESSION AL577330
 VERSION AL577330.1 GI:12940354
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 909)
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
 Location/Qualifiers
 1. 909
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS0D1085F09"
 /clone_lib="LTI_NFL006.PL2"
 /cbase_type="placenta"
 /note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT)-primer. Five prime and enriched, double-stranded cDNA was digested with NotI and cloned into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact: Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax: (1) 301 610 8371
 Email: fliang@life.com URL:
 http://fulllength.invitrogen.com"

BASE COUNT 331 a 177 c 136 g 252 t 13 others
 ORIGIN

Query Match 35.5%; Score 828; DB 9; Length 909;
 Best Local Similarity 95.2%; Pred. No. 1.3e-207;
 Matches 861; Conservative 12; Mismatches 27; Indels 4; Gaps 2;

QY 1334 ACTTGCTGCGAAGCTTCTCGAATGATCACTCACTGATCCCAACCGTCAAGTG 1453
 Db 909 ACTTGCTGCGAAGCTTCTCGAATGATCACTCACTGATCCCAACCGTCAAGTG 850
 QY 1454 TAACCTGCTGCTTTCGAAAGAAAGTTGATCTGAATTCCTGTAAGTTAGCTATAC 1513
 Db 849 TAACCTGCTGCTTTCGAAAGAAAGTTGATCTGAATTCCTGTAAGTTAGCTATAC 790
 QY 1514 AATTGAGAAATAGACGTTCACTGCGAATTTAGTGGTGAGAAATTTAGTTAGGT 1573
 Db 789 AATTGAGAAATAGACGTTCACTGCGAATTTAGTGGTGAGAAATTTAGTTAGGT 730
 QY 1574 GTTGGGATCGAAGCTCACTGCTGTTGTTCTTTATGTTGTTCTTAACATGAT 1633
 Db 729 GTTGGGATCGAAGCTCACTGCTGTTGTTCTTTATGTTGTTCTTAACATGAT 670
 QY 1634 CATAGCCAAAATTTTTCGAAAGTGTGGTGAATGTTGTTCTTTACCCACGA 1693

Db 669 CATAGCCAAAATTTTTCGAAAGTGTGGTGAATGTTGTTCTTTTACCCACGA 610
 QY 1694 AGACATCAAGATACCTGTAATTAAGCTGATAGCATATATTCATACCTGTTAGACT 1753
 Db 609 AGACATCAAGATACCTGTAATTAAGCTGATAGCATATATTCATACCTGTTAGACT 550
 QY 1754 TGGGTGAAAAGTATGAGCAGTGGAGACTAAGTGTATTAACCTTACCTGTGAATCATANGT 1813
 Db 549 TGGGTGAAAAGTATGAGCAGTGGAGACTAAGTGTATTAACCTTACCTGTGAATCATANGT 490
 QY 1814 TGTAGAAAAGCTGTCTCCATAGCTCAACAGACTGAATTCAAAGCATGTCAAGTGATA 1873
 Db 489 TGTAGAAAAGCTGTCTCCATAGCTCAACAGACTGAATTCAAAGCATGTCAAGTGATA 430
 QY 1874 GTAGATCTGTGGCATATAGAGAGGATGAGTGCCTTCCCATTCATTCTGATGAAAT 1933
 Db 429 GTAGATCTGTGGCATATAGAGAGGATGAGTGCCTTCCCATTCATTCTGATGAAAT 370
 QY 1934 TGTATATCTAGTTAAACATTTGTAATTTTCTAGTTGTATGTAAGTCTGTAAT 1993
 Db 369 TGTATATCTAGTTAAACATTTGTAATTTTCTAGTTGTATGTAAGTCTGTAAT 310
 QY 1994 AGGTATATATTTTGGCTTACATACCGTAACAATGTTGTATTTGAAATTAAT 2053
 Db 309 AGGTATATATTTTGGCTTACATACCGTAACAATGTTGTATTTGAAATTAAT 250
 QY 2054 GCCAAGTAC--AATGATGCTTTGGAATTTGGAAGATGCTTTATCTTTGAGAAGA 2111
 Db 249 GCCAAGTACATATGATGATGCTTTGGAATTTGGAAGATGCTTTATTTATTAAGA 190
 QY 2112 AATATGTTGATTAATATGTTGATGATGTC--ATATCAAGAAATGATGAAGCTTCT 2169
 Db 189 AATATGTTGATTAATATGTTGATGATGTC--ATATCAAGAAATGATGAAGCTTCT 130
 QY 2170 AAACCTGTTTACCGTACTGTGAAGGAGCGGTTTGGAGAGACCATTCATGCTGCT 2229
 Db 129 AAACCTGTTTACCGTACTGTGAAGGAGCGGTTTGGAGAGACCATTCATGCTGCT 70
 QY 2230 GTCCAGTGTCTTCTTTGTAAGTGTCTTTAACTGAGAGGCTAACCTCAAAATCTTTT 2289
 Db 69 GTCCAGTGTCTTCTTTGTAAGTGTCTTTAACTGAGAGGCTAACCTCAAAATCTTTT 10
 QY 2290 TTA 2293
 Db 9 CTGA 6

RESULT 13
 AL544755 885 bp mRNA linear EST 16-FEB-2001
 LOCUS AL544755 LTI_NFL006.PL2 Homo sapiens cDNA clone CS0D1022YJ23.5
 DEFINITION prime, mRNA sequence.
 ACCESSION AL544755
 VERSION AL544755.1 GI:12877235
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 885)
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
 Location/Qualifiers
 1. 885
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS0D1022YJ23"
 /clone_lib="LTI_NFL006.PL2"

/notes="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA
 enriched, double-stranded cDNA was digested with Not I and
 cloned into the Not I and Eco RV sites of the pCMVSPORT 6
 vector. Library was normalized. Library was constructed by
 Life Technologies. Contact : Feng Liang Life Technologies,
 a division of Invitrogen 9800 Medical Center Drive
 Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
 Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com"

BASE COUNT 247 a 211 c 228 g 198 t 1 others
 ORIGIN

Query Match 35.4%; Score 825; DB 9; Length 885;
 Best Local Similarity 99.2%; Pred. No. 7.8e-207;
 Matches 880; Conservative 1; Mismatches 1; Indels 5; Gaps 5;

QY 95 AGGGGCTTGAACGAGGCGGCTGTAGCGCTTTCACTGTAAAGCAAGATGCAATTTAG 154
 DB 1 AGGGGCTTGAACGAGGCGGCTGTAGCGCTTTCACTGTAAAGCAAGATGCAATTTAG 60
 QY 155 AAATTTTAACTACGATTTAGTCCCTGATTCCTGTGGCAACAGTGATCTTCAG 214
 DB 61 AAATTTTAACTACGATTTAGTCCCTGATTCCTGTGGCAACAGTGATCTTCAG 120
 QY 215 CGAAAGTGAACGAGGCGGCAATTTGAGTCCCTTTTAGGAGCTATGCAAGGACATC 274
 DB 121 CGAAAGTGAACGAGGCGGCAATTTGAGTCCCTTTTAGGAGCTATGCAAGGACATC 180
 QY 275 CTTTCAGTATTTAAAGAGCTTCAACAGAGTCAGATAAATCTAGCAACCCCTTCTCCG 334
 DB 181 CTTTCAGTATTTAAAGAGCTTCAACAGAGTCAGATAAATCTAGCAACCCCTTCTCCG 240
 QY 335 AGCAGATCCAGGCTCCAGTGCATAGATGAGTTTCTGGAAAGGAATGAAGTTATA 394
 DB 241 AGCAGATCCAGGCTCCAGTGCATAGATGAGTTTCTGGAAAGGAATGAAGTTATA 300
 QY 395 TTTTGTCTCAGACCTTACATAGGAGCTACACCTCGCTCCGCAATTCAGACAGCA 454
 DB 301 TTTTGTCTCAGACCTTACATAGGAGCTACACCTCGCTCCGCAATTCAGACAGCA 360
 QY 455 GTTTTCTGATCTCCCTCCGCTTCCGCCAGTGGGATGAAACAAAGTGGAAGATCGAC 514
 DB 361 GTTTTCTGATCTCCCTCCGCTTCCGCCAGTGGGATGAAACAAAGTGGAAGATCGAC 420
 QY 515 CCCAGTCATAAATATGATCTCTTATATGTCATCTCAAGCTGGGCGGCGGGAAGTA 574
 DB 421 CCCAGTCATAAATATGATCTCTTATATGTCATCTCAAGCTGGGCGGCGGGAAGTA 480
 QY 575 TGAATTCGACGAGCTGACACCACTCCCGAGCGTGGTCCATGTATGTGAGAGTGA 634
 DB 481 TGAATTCGACGAGCTGACACCACTCCCGAGCGTGGTCCATGTATGTGAGAGTGA 540
 QY 635 TCAAGAGAGGAGGAGAGAGGAGGAAATGGAAGAAATGAGAGACCTTAAGCC 693
 DB 541 TCAAGAGAGGAGGAGAGAGGAGGAAATGGAAGAAATGAGAGACCTTAAGCC 600
 QY 694 TCAGACAGGAGGCGGAGTACACCCGATCCACCTCAGCTGMACTGGCAGCGACGAG 753
 DB 601 TCAGACAGGAGGCGGAGTACACCCGATCCACCTCAGCTGMACTGGCAGCGACGAG 660
 QY 754 GACGATTCCTCAATATATCTACGGGAGGAATCTTTTACTGTGGAGGTGGCTGCTACGA 813
 DB 661 GACGATTCCTCAATATATCTACGGGAGGAATCTTTTACTGTGGAGGTGGCTGCTACGA 720
 QY 814 CTTTCTCGAGGTGGCAGCGGAGATCGGGGTGGCAGGAATCCAGTTCATGTGCTCAGA 873
 DB 721 CTTTCTCGAGGTGGCAGCGGAGATCGGGGTGGCAGGAATCCAGTTCATGTGCTCAGA 780
 QY 874 AGGAATCAAGGCGGTGTCCTTCTTAATGCTGCAACACAGTACTGTTTCATGGCAC 933
 DB 781 AGGAATCAAGGCGGTGTCCTTCTTAATGCTGCAACACAGTACTGTTTCATGGCAC 838

QY 934 CCGGAATGACTTGGGCAATCACTAGTTTGTGATCGCACAG 980
 DB 839 CCGGAATGACTTGGG-CAATCACTAG-TTGTGGTATCGCACAG 883
 RESULT 14
 AL566338 849 bp mRNA linear EST 16-FEB-2001
 LOCUS AL566338 LTI_FL013_FBrnl Homo sapiens cDNA clone CS0DF015F21 3
 DEFINITION prime, mRNA sequence.
 ACCESSION AL566338
 VERSION AL566338.1 GI:12918597
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 849)
 AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
 TITLE Full-length cDNA libraries and Polayes, D.
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: segr@genoscope.cns.fr, Web : www.genoscope.cns.fr.
 Location/Qualifiers
 1. 849
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS0DF015F21"
 /clone_lib="LTI_FL013_FBrnl"
 /dev_stage="pooled tissue from post conception fetuses (20
 week, 24 week and 26 week)"
 /lab_host="DH10B"
 /note="Organ: Fetal brain; Vector: pCMVSPORT 6; 1st strand
 cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-stranded cDNA was digested with Not I
 and cloned into the Not I and Eco RV sites of the
 pCMVSPORT 6 vector. Library was constructed by Life
 Technologies. Contact : Feng Liang Life Technologies, a
 division of Invitrogen 9800 Medical Center Drive Rockville
 , Maryland 20850, USA Fax : (1) 301 610 8371 Email :
 fliang@lifetech.com URL :
 http://fulllength.invitrogen.com"

BASE COUNT 309 a 175 c 120 g 235 t 10 others
 ORIGIN

Query Match 34.7%; Score 808.6; DB 9; Length 849;
 Best Local Similarity 97.5%; Pred. No. 1.7e-202;
 Matches 830; Conservative 10; Mismatches 9; Indels 2; Gaps 2;

QY 1450 CGTGTAACTTGTGCTTTTGCAGAAAGAGTTGATCTGAAATTCCTCTGTAGAAATTTAGCTT 1509
 DB 849 CGTGTAACTTGTGCTTTTGCAGAAAGAGTTGATCTGAAATTCCTCTGTAGAAATTTAGCTT 790
 QY 1510 ATACAATTCAGAGAAATAGCAGTTTCACTGCCACTTTTGTAGTGGTGAGAAATTTAGTTT 1569
 DB 789 ATACAATTCAGAGAAATWASAG-TTCACTGCCACTTTTGTAGTGGTGAGAAATTTAGTTT 731
 QY 1570 AGTGTGTGGATTCGACCTCAGTTTCTGTGTTCTTTTATGTGGTGGTCTTATACAT 1629
 DB 730 AGTGTGTGGATTCGACCTCAGTTTCTGTGTTCTTTTATGTGGTGGTCTTATACAT 671
 QY 1630 GAATCATAGCAAAACTTTTGGAACTGTGTTGAGATAGTGGTCTTTTACCCC 1689
 DB 670 GAWTCATAGCAAAACTTTTGGAACTGTGTTGAGATAGTGGTCTTTTACCCC 611
 QY 1690 ACCAGACATCAAGATACACTTTGTAAATAAGCTCATAGCATATATTCATACCTGTTGA 1749
 DB 610 ACCAGACATCAAGATACACTTTGTAAATAAGCTCATAGCATATATTCATACCTGTTGA 551
 QY 1750 CACTTGGGTGAAAGATATGGCAGTGGGAGACTAAGATGATTTAACTACCTGTTGAATCAT 1809

```

|||||
Db 550 CACTGGTGAAGATATGCGAGTGGAGACATGATATTAACCTACCTGTAATCAT 491
QY 1810 ATGTTGTGAGAAAGCTCTCCCATGCTAACAGACCTGAATCAAGCATGCAAGTG 1869
Db 490 ATGTTGTGAGAAAGCTCTCCCATGCTAACAGACCTGAATCAAGCATGCAAGTG 431
QY 1870 GATAGTATGCTGTGGCGCATATGAGAGGATGCACTGCTTCCCATCATCTCTGATG 1929
Db 430 CATAGTATGCTGTGGCGCATATGAGAGGATGCACTGCTTCCCATCATCTCTGATG 371
QY 1930 GAATGTTATCTAGCTTAACATTTGTAATTTTCTTCTAGTGTATGATGCTGCT 1989
Db 370 GAATGTTATCTAGCTTAACATTTGTAATTTTCTTCTAGTGTATGATGCTGCT 311
QY 1990 AATATGCTTATATTTTGGCCTTACATACCGTACATGTTTGTCAATTTTGAATACT 2049
Db 310 AATATGCTTATATTTTGGCCTTACATACCGTACATGTTTGTCAATTTTGAATACT 251
QY 2050 TAATGCCAATGTAACAATGATGCTTTGAAATTTGGAAATGTTTATCTTTGAGAG 2109
Db 250 TAATGCCAATGTAACAATGATGCTTTGAAATTTGGAAATGTTTATCTTTGAGAG 191
QY 2110 CAATATGTTGATTAATAGCTTTGATGTTGATGATGATGATGATGATGATGATG 2169
Db 190 CAATATGTTGATTAATAGCTTTGATGTTGATGATGATGATGATGATGATGATG 131
QY 2170 AAACCTGTTACGCTACTTGTGTAAGAGGAGCGGTTTGGAGAGACATGTCATCGCT 2229
Db 130 AAACCTGTTACGCTACTTGTGTAAGAGGAGCGGTTTGGAGAGACATGTCATCGCT 72
QY 2230 GTCCAACTGTTCTTCTTTAAAGTCTTTAACTGGAGAGGCTAACCTCAAAATACTTTT 2289
Db 71 GTCCAACTGTTCTTCTTTAAAGTCTTTAACTGGAGAGGCTAACCTCAAAATACTTTT 12
QY 2290 TTACTGCTATT 2300
Db 11 TTACTGCTATT 1

RESULT 15
BG695987 856 bp mRNA linear EST 07-MAY-2001
LOCUS BG695987
DEFINITION BG695987.1 NC_1_CGAP_Skn3 Homo sapiens cDNA clone IMAGE:480918 5',
            mRNA sequence.
ACCESSION BG695987
VERSION BG695987.1 GI:13960662
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 856)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
            Email: cga@bbs-remail.nih.gov.
            Tissue Procurement: James Cleaver, M.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA
            Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNL at:
            http://image.jhmi.gov
            Plate: LLM10692 row: 1 column: 07
            High quality sequence stop: 821.
            Location/Qualifiers
                1..856
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:480918"
                /clone_1ib="NCI_CGAP_Skn3"

```

```

/lab host="DH10B (T1 phage-resistant)"
/Note="Organ: skin; Vector: pCMV-Sport6; Site_1: NotI;
       Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
       Average insert size 1.5kb. Library constructed by Life
       Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT      230 a      148 c      174 g      304 t
ORIGIN
Query Match      34.6%; Score 806.6; DB 12; Length 856;
Best Local Similarity 98.6%; Pred. No. 5.7e-202;
Matches 845; Conservative 0; Mismatches 9; Indels 3; Gaps 3;
QY 1208 GTAGCGTTCTCTTCTCCCTCATGCTTCTCATGTTTGGCATATATATCTAGTATCA 1267
Db 1 GTAGCGTTCTCTTCTCCCTCATGCTTCTCATGTTTGGCATATATATCTAGTATCA 60
QY 1268 AGACTAACCTTTGTTGATATATTAAGTATACACCGTTGTTTATCATCTTTGGGAAC 1327
Db 61 AGACTAACCTTTGTTGATATATTAAGTATACACCGTTGTTTATCATCTTTGGGAAC 120
QY 1328 CAGGAAAGCGTTTGGAAAGAGATACCTTTTCCAGATTCGAGATTCCTGACTCTTTCG 1387
Db 121 CAGGAAAGCGTTTGGAAAGAGATACCTTTTCCAGATTCGAGATTCCTGACTCTTTCG 180
QY 1388 AACGCACTTGTCTGGGAACTCTTCTGGAATGCAATGCTCAGCATCCCAACCGTGC 1447
Db 181 AACGCACTTGTCTGGGAACTCTTCTGGAATGCAATGCTCAGCATCCCAACCGTGC 240
QY 1448 AACGTTAACTTGTGCTTTTGGCAAAAGAGTGAATCTGATGAAATTTAGC 1507
Db 241 AACGTTAACTTGTGCTTTTGGCAAAAGAGTGAATCTGATGAAATTTAGC 300
QY 1508 TTATACAAATTCAGAAATAGAGATTTGACGCGCACTTTAGGGGGAGAAATTTAGT 1567
Db 301 TTATACAAATTCAGAAATAGAGATTTGACGCGCACTTTAGGGGGAGAAATTTAGT 360
QY 1568 TTAGGTGTTTGGGATCGGACCTCAGTTTCTGTTCTTTTATGNGSGTTCATAC 1627
Db 361 TTAGGTGTTTGGGATCGGACCTCAGTTTCTGTTCTTTTATGNGSGTTCATAC 420
QY 1628 ATGAATCATAGCCAAAACTTTTTCGAACTGTTGTTGATGATGTTGTTCTTTAC 1687
Db 421 ATGAATCATAGCCAAAACTTTTTCGAACTGTTGTTGATGATGTTGTTCTTTAC 480
QY 1688 CCAGCAAGACATCAAGATACCTGTTAAATTAAGCTGATACATATATTCATCTGTG 1747
Db 481 CCAGCAAGACATCAAGATACCTGTTAAATTAAGCTGATACATATATTCATCTGTG 540
QY 1748 TACACTTGGGTGAAAAGATATGCGATGAGAGCTAAGATGATTNACTTACCTGTGAATC 1807
Db 541 TACACTTGGGTGAAAAGATATGCGATGAGAGCTAAGATGATTNACTTACCTGTGAATC 600
QY 1808 ATATGTTGTGAGAAA-GCTGTTTCCCATGCTTAACAGACTGTAATCAAGCATGCA 1866
Db 601 ATATGTTGTGAGAAAAGCTGTTTCCCATGCTTAACAGACTGTAATCAAGCATGCA 660
QY 1867 GTGGATATGATCTGTGGCATATGAGAGGATGAGGATGAGGATGAGGATGAGGATGAG 1926
Db 661 GTGGATATGATCTGTGGCATATGAGAGGATGAGGATGAGGATGAGGATGAGGATGAG 720
QY 1927 ATGAAATGTTATATAGTATGA- TTGTAATTTTCTTCTAGTGTATATGATGATG 1985
Db 721 ATGAAATGTTATATAGTATGA- TTGTAATTTTCTTCTAGTGTATATGATGATG 780
QY 1986 TGTGAAATAGTATATATTTTGGCTTCAATATCCGTAACATGTTTGTCAATTTGAAA 2045
Db 781 TGTGAAATAGTATATATTTTGGCTTCAATATCCGTAACATGTTTGTCAATTTGAAA 839
QY 2046 TACTTAATGCCAAGTA 2062
Db 840 TACTTAATGCCAAGTA 856

```

Sun Dec 15 08:38:04 2002

Search completed: December 14, 2002, 18:11:18
Job time : 2612.11 secs

us-09-782-953-11.rst

THIS PAGE BLANK (USPRO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2002, 14:10:15 ; Search time 97.497 Seconds
(without alignments)
13789.580 Million cell updates/sec

Title: US-09-782-953-5

Perfect score: 597

Sequence: 1 atggatttttaggaacttag.....cacgatccaccttagctga 597

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq 101002.*

```

1: /SID2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	597	100.0	597	24	AAD30153
2	515.8	86.4	597	24	AAD30152
3	504.4	84.5	599	24	AAD30151
4	470.6	78.8	2331	24	AAD30155
5	459.6	77.0	2348	24	AAS94805
6	408.2	68.4	2212	24	AAD30156
7	399.6	66.9	2173	24	ABN96843
8	398	66.7	2174	20	AAK01282
9	395.4	66.2	2358	21	AAF18328

10	273.8	45.9	442	21	AAC01774	Human secreted pro
11	214.2	35.9	3184	24	AAD30157	Human MCIP associa
12	214.2	35.9	3184	24	ABL61768	Colon adenocarcino
13	214.2	35.9	3184	24	ABL65187	Lung cancer relate
14	209.6	35.1	934	22	AAF25338	Nucleotide sequenc
15	209.6	35.1	3159	24	ABA91457	Human Down syndrom
16	201.4	33.7	594	24	AAD30154	Mouse MCIP associa
17	191.2	32.0	720	24	AAD30159	Human MCIP associa
18	191.2	32.0	828	24	AAD30158	Human MCIP associa
19	186.2	31.2	615	24	ABA91463	Rat Down syndrome
20	155	26.0	1021	24	ABA91463	Rat Down syndrome
21	148	24.8	640	24	ABQ59719	Human colon cancer
22	144.6	24.2	412	22	ABA56172	Human foetal liver
23	144.6	24.2	412	22	AAK04364	Human brain expres
24	144.6	24.2	412	22	AAI14449	Probe #4382 for ge
25	144.6	24.2	412	22	AAI15821	Probe #4507 used t
26	144.6	24.2	412	22	AAI04272	Probe #4263 used t
27	144.6	24.2	412	22	ABQ44119	Human genome-deriv
28	144.6	24.2	446	22	ABA43406	Human breast cell
29	144.6	24.2	446	22	ABA53852	Probe #2064 for ge
30	144.6	24.2	446	22	ABA23598	Human brain expres
31	144.6	24.2	446	22	AAK02113	Human bone marrow
32	144.6	24.2	446	22	AAK27563	Probe #2081 for ge
33	144.6	24.2	446	22	AAI12148	Probe #2179 used t
34	144.6	24.2	446	22	AAI33493	Probe #2052 used t
35	144.6	24.2	446	22	AAI02061	Human genome-deriv
36	144.6	24.2	446	24	ABQ42132	Human breast cell
37	123	20.6	486	22	ABA52554	Human foetal liver
38	123	20.6	486	22	ABA22343	Probe #809 for gen
39	123	20.6	486	22	AAK00816	Human brain expres
40	123	20.6	486	22	AAK26270	Human bone marrow
41	123	20.6	486	22	AAI10903	Probe #836 for gen
42	123	20.6	486	22	AAI32163	Probe #849 used to
43	123	20.6	486	22	AAI00825	Probe #816 used to
44	123	20.6	486	22	ABS00858	Human genome-deriv
45	123	20.6	486	24	ABS00858	Human genome-deriv

ALIGNMENTS

RESULT 1
AAD30153
ID AAD30153 standard; DNA; 597 BP.
XX AAD30153;
AC AAD30153;
XX

17-MAY-2002 (first entry)
XX

Mouse MCIP associated DNA #3.
XX

Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;
heart failure; cardiomyopathy; heart disease; mouse; gene; ds.
XX

Mus musculus.
XX

Key Location/Qualifiers
CDS 1..597
FT /tag= a
FT /product= "Mouse MCIP associated protein #3"

WO200204491-A2.
XX

17-JAN-2002.
XX

06-JUL-2001; 2001WO-US21662.
XX

07-JUL-2000; 2000US-216601P.
PR

13-FEB-2001; 2001US-0782953.
XX

(TEXA) UNIV TEXAS SYSTEM.
PA

(WILL) WILLIAMS S R.
PA

(ROTH) ROTHERMEL B.
PA

XX Williams SR, Rothermel B;
 XX WPI; 2002-179698/23.
 DR P-PSDB; AAE18912.
 XX
 XX Screening for modulators of muscle calcineurin interacting protein
 PT (MCIP) binding, expression or phosphorylation, useful for treating
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,
 PT calcineurin and a test compound -
 XX
 PS Disclosure; Page 150-151; 174pp; English.
 XX
 XX The invention relates to muscle calcineurin interacting proteins (MCIPs)
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical
 CC complex with the catalytic subunit of calcineurin and increased levels
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate
 CC transcription of certain target genes. The invention also relates to
 CC methods for identifying modulators of MCIP binding, expression or
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin
 CC may be used for treating cardiac hypertrophy and heart failure.
 CC Antibodies to MCIP can be used in characterizing the MCIP content of
 CC healthy and diseased tissues and subsequently for determining the
 CC presence or absence of cardiomyopathy or as predictor of heart disease.
 CC The present sequence is mouse MCIP associated DNA.
 CC Note: This sequence has been described as human MCIP3 encoding DNA
 CC in the specification, however the sequence seems to be a polynucleotide
 CC encoding a MCIP associated protein.
 XX
 XX Sequence 597 BP; 170 A; 156 C; 142 G; 129 T; 0 other;
 XX
 XX Query Match 100.0%; Score 597; DB 24; Length 597;
 XX Best Local Similarity 100.0%; Pred. No. 5,3e-176; Indels 0; Gaps 0;
 XX Matches 597; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX 1 ATGGATTTTGGGAGCTTAGTACATTTTATGCTCCCTGATTCCTTGTGGCAACGAT 60
 XX 1 ATGGATTTTGGGAGCTTAGTACATTTTATGCTCCCTGATTCCTTGTGGCAACGAT 60
 XX
 XX 61 GATGCTCTCAGCGAAGTAGAGACGAGGCCAATTGATCCCTCTTCAGAAATGAC 120
 XX 61 GATGCTCTCAGCGAAGTAGAGACGAGGCCAATTGATCCCTCTTCAGAAATGAC 120
 XX
 XX 121 AAGGACACCACTTCCGATATTTTAAAGAGCTTCAACGTCGCGATTAACCTTCAGCAAC 180
 XX 121 AAGGACACCACTTCCGATATTTTAAAGAGCTTCAACGTCGCGATTAACCTTCAGCAAC 180
 XX
 XX 181 CCTTATCTGACGCGATGCGAGGCTGCGCTGCACAAAGCCGAGTTCTGGGGAAGAA 240
 XX 181 CCTTATCTGACGCGATGCGAGGCTGCGCTGCACAAAGCCGAGTTCTGGGGAAGAA 240
 XX
 XX 241 ATGAAGTTGATTTTGTCTCAGCTTACACATAGGAATTCACACTGCTCCGCCCAAT 300
 XX 241 ATGAAGTTGATTTTGTCTCAGCTTACACATAGGAATTCACACTGCTCCGCCCAAT 300
 XX
 XX 301 CCCGCAAAAGTCTCTCATCTCCCTCCGCGCTCTCTCCGCTGCTGGAACAAGTA 360
 XX 301 CCCGCAAAAGTCTCTCATCTCCCTCCGCGCTCTCTCCGCTGCTGGAACAAGTA 360
 XX
 XX 361 GAAGATGCCACCCCGCTCATTAATTAATTCATTTTATATGATCCCAAGTGGGGCCA 420
 XX 361 GAAGATGCCACCCCGCTCATTAATTAATTCATTTTATATGATCCCAAGTGGGGCCA 420
 XX
 XX 421 GGAAGAAGATTAAGTATGATGATGAGAGAGACCCCACTCCAGTGTGTGTCACGTTG 480
 XX 421 GGAAGAAGATTAAGTATGATGATGAGAGAGACCCCACTCCAGTGTGTGTCACGTTG 480
 XX
 XX 481 TGTGAGAGTACCAAGAGATGAGAGAGAGAGAGAGAGATGAGAGAGAGAGAGAGAGAG 540
 XX 481 TGTGAGAGTACCAAGAGATGAGAGAGAGAGAGAGAGATGAGAGAGAGAGAGAGAGAG 540
 XX
 XX 541 AAGCCCAAAATCATCCGAGACAGGAGACGAGATACACACCATTCACCTTAGCTGA 597
 XX 541 AAGCCCAAAATCATCCGAGACAGGAGACGAGATACACACCATTCACCTTAGCTGA 597

DB 541 AAGCCCAAAATCATCCGAGACAGGAGACGAGATACACACCATTCACCTTAGCTGA 597
 XX
 XX RESULT 2
 XX AAD30152
 XX ID AAD30152 standard; DNA; 597 BP.
 XX
 XX AAD30152;
 XX
 XX 17-MAY-2002 (first entry)
 XX
 XX Human MCIP associated DNA #2.
 XX
 XX Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;
 XX heart failure; cardiomyopathy; heart disease; human; gene; ds.
 XX
 XX Homo sapiens.
 XX
 XX Key Location/Qualifiers
 XX CDS 1..597
 XX FT /*tag= a
 XX FT /product= "Human MCIP associated protein #2"
 XX
 XX MO200204491-A2.
 XX
 XX 17-JUN-2002.
 XX
 XX 06-JUL-2001; 2001WO-US21662.
 XX
 XX 07-JUL-2000; 2000US-216601P.
 XX
 XX 13-FEB-2001; 2001US-0782953.
 XX
 XX (TEXA) UNIV TEXAS SYSTEM.
 XX (WILL/) WILLIAMS S R.
 XX (ROTH/) ROTHERMEL B.
 XX
 XX Williams SR, Rothermel B;
 XX WPI; 2002-179698/23.
 XX P-PSDB; AAE18911.
 XX
 XX Screening for modulators of muscle calcineurin interacting protein
 PT (MCIP) binding, expression or phosphorylation, useful for treating
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,
 PT calcineurin and a test compound -
 XX
 PS Claim 95; Page 147-148; 174pp; English.
 XX
 XX The invention relates to muscle calcineurin interacting proteins (MCIPs)
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical
 CC complex with the catalytic subunit of calcineurin and increased levels
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate
 CC transcription of certain target genes. The invention also relates to
 CC methods for identifying modulators of MCIP binding, expression or
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin
 CC may be used for treating cardiac hypertrophy and heart failure.
 CC Antibodies to MCIP can be used in characterizing the MCIP content of
 CC healthy and diseased tissues and subsequently for determining the
 CC presence or absence of cardiomyopathy or as predictor of heart disease.
 CC The present sequence is human MCIP associated DNA.
 XX
 XX Sequence 597 BP; 160 A; 170 C; 151 G; 116 T; 0 other;
 XX
 XX Query Match 86.4%; Score 515.8; DB 24; Length 597;
 XX Best Local Similarity 94.4%; Pred. No. 1.2e-150; Indels 0; Gaps 0;
 XX Matches 535; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
 XX
 XX 31 AGCTCCCTGATTCCTTGTGTGCGAAGCATGATGCTTCAGCGAAGTAGAGACGAGGCC 90
 XX 31 AGCTCCCTGATTCCTTGTGTGCGAAGCATGATGCTTCAGCGAAGTAGAGACGAGGCC 90
 XX
 XX 31 AGCGCCACCATTCCTGCGACCTGAGCCCGCGGTGTTGTGAGAGCGCTGTGCGGGGCC 90
 XX 31 AGCGCCACCATTCCTGCGACCTGAGCCCGCGGTGTTGTGAGAGCGCTGTGCGGGGCC 90
 XX
 XX 91 AAATTTGAATCCCTCTTCAGAAATATGACAAAGACACCACTTCGATTTTAAAGAC 150
 XX 91 AAATTTGAATCCCTCTTCAGAAATATGACAAAGACACCACTTCGATTTTAAAGAC 150

Db 91 "AAATTGAAATCCCTCTTCAGAACATATGACAAAGGACACCACTTCCAGTATTTTAAAGAGC 150
 Qy 151 TTCAACAGTGTCCGATAACTTACGAAACCCCTTATCTGCAGCGATGCCAGGCTGGG 210
 Db 151 TTCAACAGTGTCCGATAACTTACGAAACCCCTTATCTGCAGCGATGCCAGGCTGGG 210
 Qy 211 CTGCAACAGCGAGTTCCTGGGAGGAAATGAAGTTGTATTTTGTCTCAGACTTTACAC 270
 Db 211 CTGCAACAGCGAGTTCCTGGGAGGAAATGAAGTTGTATTTTGTCTCAGACTTTACAC 270
 Qy 271 ATAGGAAGTTCACACTCGCTCCGCCCAATCCGACAAACAGTTCTCATCTCCCTCCG 330
 Db 271 ATAGGAAGTTCACACTCGCTCCGCCCAATCCGACAAACAGTTCTCATCTCCCTCCG 330
 Qy 331 GCCTCTCTCCGTTGGTGGAAACAAGTAGAGATGCAACCCCGTCATAAAATTACGAT 390
 Db 331 GCCTCTCTCCGTTGGTGGAAACAAGTAGAGATGCAACCCCGTCATAAAATTACGAT 390
 Qy 391 CTTTATATGCAATCCAGCTGGGCGCAGGAGAGATGATGAATCTGATGCAAGCGACA 450
 Db 391 CTTTATATGCAATCCAGCTGGGCGCAGGAGAGATGATGAATCTGATGCAAGCGACA 450
 Qy 451 GACCCCACTCCCACTGGTGGTCCAGCTGTGTCAGAGTGACCAAGAGAAATGAGGAGAA 510
 Db 451 GACCCCACTCCCACTGGTGGTCCAGCTGTGTCAGAGTGACCAAGAGAAATGAGGAGAA 510
 Qy 511 GAGGAAGATGAGAGAAATGAAGAGACCCCAAGCCCAAAATCATCCAGACAGGAGACCG 570
 Db 511 GAGGAAGATGAGAGAAATGAAGAGACCCCAAGCCCAAAATCATCCAGACAGGAGACCG 570
 Qy 571 GAGTACACACCGATCCACCTTAGCTGA 597
 Db 571 GAGTACACACCGATCCACCTTAGCTGA 597

RESULT 3

AAD30151
 ID AAD30151 standard; DNA; 599 Bp.
 XX
 AC AAD30151;
 XX
 DT 17-MAY-2002 (first entry)
 XX
 DE Human MCIP associated DNA #1.
 XX
 KW Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;
 KW heart failure; cardiomyopathy; heart disease; human; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200204491-A2.
 XX
 PD 17-JAN-2002.
 XX
 PF 06-JUL-2001; 2001WO-US21662.
 XX
 PR 07-JUL-2000; 2000US-216601P.
 PR 13-FEB-2001; 2001US-0782953.
 XX
 PA (TEXA) UNIV TEXAS SYSTEM.
 PA (WILL/) WILLIAMS S R.
 PA (ROTH/) ROTHERMEL B.
 XX
 PI Williams SR, Rothermel B;
 XX
 DR WPI; 2002-179698/23.
 XX
 PT Screening for modulators of muscle calcineurin interacting protein
 PT (MCIP) binding, expression or phosphorylation, useful for treating
 PT cardiac hypertrophy or heart failure, comprises mixing MCIP,
 XX calcineurin and a test compound -
 PS Disclosure; Page 147; 174pp; English.

XX The invention relates to muscle calcineurin interacting proteins (MCIPs)
 CC and nucleic acid molecules encoding such proteins. MCIPs form a physical
 CC complex with the catalytic subunit of calcineurin and increased levels
 CC of MCIPs correspond to a reduced ability of calcineurin to stimulate
 CC transcription of certain target genes. The invention also relates to
 CC methods for identifying modulators of MCIP binding, expression or
 CC phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin
 CC may be used for treating cardiac hypertrophy and heart failure.
 CC Antibodies to MCIP can be used in characterizing the MCIP content of
 CC healthy and diseased tissues and subsequently for determining the
 CC presence or absence of cardiomyopathy or as predictor of heart disease.
 CC The present sequence is human MCIP associated DNA.

SQ Sequence 599 Bp; 164 A; 150 C; 149 G; 136 T; 0 other;

Query Match 84.5%; Score 504.4; DB 24; Length 599;

Best Local Similarity 99.6%; Pred. No. 4.3e-147;

Matches 516; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 4 GATTTTAGGAGCTTTAGCTAGCAATTTAGCTCCCTGATTTGTTGGCAACGATGAT 63
 Db 83 GATTTTAGGAGCTTTAGCTAGCAATTTAGCTCCCTGATTTGTTGGCAACGATGAT 142
 Qy 64 GTCTTCAGCGAAAGTGAGACAGGCGCAAAATTTGAATCCCTCTTCAGACATATCACAAG 123
 Db 143 GTCTTCAGCGAAAGTGAGACAGGCGCAAAATTTGAATCCCTCTTCAGACATATCACAAG 202
 Qy 124 GACACACCTTCCAGTATTTTAAAGAGCTTCAAACTGTCGGGATAAACTTCAGCAACCCC 183
 Db 203 GACACACCTTCCAGTATTTTAAAGAGCTTCAAACTGTCGGGATAAACTTCAGCAACCCC 262
 Qy 184 TTATCTGCAGCGATGCCAGGCTGCGGCTGCACAAGACCGAGTTCTGGGGAAGAAATG 243
 Db 263 TTATCTGCAGCGATGCCAGGCTGCGGCTGCACAAGACCGAGTTCTGGGGAAGAAATG 322
 Qy 244 AAGTTGTATTTTGTCTCAGACTTTACACATAGCAAGTTCCACACCTGGCTCCGCCAATCCC 303
 Db 323 AAGTTGTATTTTGTCTCAGACTTTACACATAGCAAGTTCCACACCTGGCTCCGCCAATCCC 381
 Qy 304 GACAAACAGTTCTCATCTCCCTCCCGGCTCTCTCCCGTTGGCTGGAAACAAGTAGAA 363
 Db 382 GACAAACAGTTCTCATCTCCCTCCCGGCTCTCTCCCGTTGGCTGGAAACAAGTAGAA 441
 Qy 364 GATGCCACCCCGTCATAAAATTCAGATCTTTATATGSCATCTCAAGCTGGGCGCAGGA 423
 Db 442 GATGCCACCCCGTCATAAAATTCAGATCTTTATATGSCATCTCAAGCTGGGCGCAGGA 501
 Qy 424 GAGAAGTATGAATGCACTGCATGCGACGACAGACCCCACTCCAGTGTGGTCCACGTGTGT 483
 Db 502 GAGAAGTATGAATGCACTGCATGCGACGACAGACCCCACTCCAGTGTGGTCCACGTGTGT 561
 Qy 484 GAGAGTGACCAAGAGATGAGGAGGAAGAGAGAT 521
 Db 562 GAGAGTGACCAAGAGATGAGGAGGAAGAGAGAT 599

RESULT 4

AAD30155
 ID AAD30155 standard; DNA; 2331 Bp.

XX

AC AAD30155;

XX

DT 17-MAY-2002 (first entry)

XX

DE Human MCIP associated DNA #1.

XX

KW Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;

KW heart failure; cardiomyopathy; heart disease; human; gene; ds.

XX

OS Homo sapiens.

XX

Key Location/Qualifiers

FT CDS 144..737
 FT /*tag= a
 FT /product= "Human MCIP associated protein #1"
 FT
 XX MO20020491-A2.
 XX 17-JAN-2002.
 XX
 XX 06-JUL-2001; 2001WO-US21662.
 XX
 XX 07-JUL-2000; 2000US-216601P.
 XX 13-FEB-2001; 2001US-0782953.
 XX
 XX (TEXA) UNIV TEXAS SYSTEM.
 XX (WILL) WILLIAMS S R.
 XX (ROTH) ROTHERMEL B.
 XX
 XX Williams SR, Rothermel B;
 XX
 XX WPI, 2002-179698/23.
 XX P-PEDB: AAE18914.
 XX
 XX Screening for modulators of muscle calcineurin interacting protein
 XX (MCIP) binding, expression or phosphorylation, useful for treating
 XX cardiac hypertrophy or heart failure, comprises mixing MCIP,
 XX calcineurin and a test compound -
 XX
 XX Claim 72; Page 155-157; 174pp; English.
 XX
 XX The invention relates to muscle calcineurin interacting proteins (MCIPs)
 XX and nucleic acid molecules encoding such proteins. MCIPs form a physical
 XX complex with the catalytic subunit of calcineurin and increased levels
 XX of MCIPs correspond to a reduced ability of calcineurin to stimulate
 XX transcription of certain target genes. The invention also relates to
 XX methods for identifying modulators of MCIP binding, expression or
 XX phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin
 XX can be used for treating cardiac hypertrophy and heart failure.
 XX Antibodies to MCIP can be used in characterizing the MCIP content of
 XX healthy and diseased tissues and subsequently for determining the
 XX presence or absence of cardiomyopathy or as predictor of heart disease.
 XX The present sequence is human MCIP associated DNA.
 XX Note: This sequence has been described as a promoter in claim 72 of
 XX the specification, however the sequence seems to be a polynucleotide
 XX encoding a MCIP associated protein.
 XX
 XX Sequence 2331 BP; 630 A; 470 C; 547 G; 684 T; 0 other;
 XX
 XX Query Match 78.8%; Score 470.6; DB 24; Length 2331;
 XX Best Local Similarity 87.9%; Pred. No. 3.1e-136;
 XX Matches 525; Conservative 0; Mismatches 69; Indels 3; Gaps 1;
 XX
 XX 1 ATGATTTTGGGACTTACTCAATTTAGCTCCCTGATGCTGTGGGAAAGCAT 60
 XX 144 ATGATTTTGGAACTTAACCTTACAGTTTAACTCCCTGATGCTGTGGGAAAGCAT 203
 XX
 XX 61 GATGCTCTTACGCGAAGTGAAGCCAGGCGCAATTTGATCCCTCTTACAGACATATGAC 120
 XX 204 GATGCTCTTACGCGAAGTGAAGCCAGGCGCAATTTGATCCCTCTTACAGACATATGAC 263
 XX
 XX 121 AAGGACACCACTTCCAGTATTTTAAAGAGCTTCAACGCTGCGGATTAACCTTACAGAAC 180
 XX 264 AAGGACATCACTTCCAGTATTTTAAAGAGCTTCAACGAGTCAAGATTAACCTTACAGAAC 323
 XX
 XX 181 CCCTTATCTGAGCGGATGCGAGGCTGCGGCTGCAACAAGCCGAGTTCCTGGGAAAGAA 240
 XX 324 CCCTTCTCCGAGAGAGTCCAGGCTTCAAGCTGATTAAGTGAAGTTCCTGGGAAAGAA 303
 XX
 XX 241 ATGAAGTTGATTTTGTCTCAGACTTTACACATAGAAAGTTCAACCTGCTCGCCCAAT 300
 XX 384 ATGAAGTTATTTTGTCTCAGACTTTACACATAGAAAGTTCAACCTGCTCGCCCAAT 443
 XX
 XX 301 CCCGACAAACGTTCTCATCTCCCTCGGCTCTCTCCCGTGGCTGGAAACAGTA 360

DB 444 CCAGACAAAGCATTTCTGATCTCCCTCCCGGCTTCCGCCAGTGGAGTGAACAAGTG 503
 QY 361 GAAGATGCGACCCCGCTCATTAATTAATTCATCTTTTATATGATCTCAAGCTGGGGCCA 420
 DB 504 GAAGATGCGACCCCGCTCATTAATTAATTCATCTTTTATATGATCTCAAGCTGGGGCCA 563
 QY 421 GAGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
 DB 564 GGGGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 623
 QY 481 TGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
 DB 624 TGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 680
 QY 541 AAGCCAAATTCATCTCAGACACGAGACCGGAGTACACACCGATCCTTACTTA 597
 DB 681 AAGCCAAATTCATCTCAGACACGAGACCGGAGTACACCGGATCCTTACTTA 737
 XX
 XX RESULT 5
 XX AAS94805
 XX ID AAS94805 standard; DNA; 2348 BP.
 XX
 XX AC AAS94805;
 XX
 XX 14-FEB-2002 (first entry)
 XX
 XX Human DNA sequence #60 expressed during foam cell differentiation.
 XX
 XX Human: foam cell differentiation; atherosclerosis; cerebral stroke;
 XX cardiovascular disorder; coronary artery disease; gene therapy; ds.
 XX
 XX Homo sapiens.
 XX
 XX MO200177389-A2.
 XX
 XX 18-OCT-2001.
 XX
 XX 04-APR-2001; 2001WO-US11128.
 XX
 XX 05-APR-2000; 2000US-195106P.
 XX
 XX (INCY-) INCYTE GENOMICS INC.
 XX
 XX Shiffman D, Somogyi R, Lawn R, Seilhamer JJ, Porter GJ, Mikita T;
 XX Tai J;
 XX
 XX WPI, 2002-010925/01.
 XX
 XX Composition useful for diagnosis of conditions, disorders or diseases
 XX PT associated with atherosclerosis, comprises several polynucleotides that
 XX are differentially expressed in foam cell development -
 XX
 XX Claim 1; Page 112-113; 315pp; English.
 XX
 XX The present invention relates to the isolation of human polynucleotide
 XX sequences that are differentially expressed during foam cell
 XX differentiation. The polynucleotide sequences of the invention or a
 XX composition comprising these polynucleotide sequences are useful as a high
 XX throughput method for detecting altered expression of one or more
 XX polynucleotides in a sample. The polynucleotides can be used in the
 XX diagnosis of disorders associated with foam cell development such as
 XX atherosclerosis, cerebral stroke, and cardiovascular disorders such as
 XX coronary artery disease. The polynucleotide sequences can also be used
 XX as PCR primers and probes. The polynucleotides of the invention are also
 XX useful in gene therapy. AAS94746-AAS95021 represent the human
 XX polynucleotide sequences of the invention which are differentially
 XX expressed during foam cell differentiation.
 XX
 XX Sequence 2348 BP; 638 A; 473 C; 550 G; 687 T; 0 other;
 XX
 XX Query Match 77.0%; Score 459.6; DB 24; Length 2348;
 XX Best Local Similarity 87.8%; Pred. No. 8.4e-133;

Matches: 525; Conservative 0; Mismatches 69; Indels 4; Gaps 2;

```

QY 1 ATGGATTTTAGGACTTTAGCTACAAATTTAGCTCCCTGATTGCTGTGGCAACCAT 60
Db 151 ATGCATTTTAAAGACTTTAACTACAGTTTAGCTCCCTGATTGCTGTGGCAACCAT 210
QY 61 GATGCTTTCAGGCAAGTGAACACAGGCGCAAAATTTGAATCCCTTTTCAAGCAATATGAC 120
Db 211 GATATCTTTCAGGCAAGTGAACACAGGCGCAAAATTTGAATCCCTTTTAGGACGTATGAC 270
QY 121 AAGGACACCACTTCCAGTATTTTAAAGAGCTTCAAACTGTCCGATGAATTTTCAAGCAAC 180
Db 271 AAGGACATCACTTTCAGTATTTTAAAGAGCTTCAAACTGTCCGATGAATTTTCAAGCAAC 330
QY 181 CCCTTATCTGAGCGATGCCAGCTGCGGCTGCACAAAGCGAGTTCTTGGGGAAGGAA 240
Db 331 CCCTTCTCCGACGAGATGCCAGCTCCAGCTGCATAGAGTGTGTTCTGGGAAGGAA 390
QY 241 ATGAAGTTTGTATTTTGTCTCAGACTTTTACATAGGAAGTTTCAACCTTGGCTCCGCCAA 299
Db 391 ATGAAGTTTATTTTGTCTCAGACTTTTACATAGGAAGTTTCAACCTTGGCTCCGCCAA 450
QY 300 TCCGACAAACAGTTCTCTCATCTCCCTCCGCTCTCTCTCCGCTGCTGGAACAGT 359
Db 451 TCCGACAAACAGTTCTCTCATCTCCCTCCGCTCTCTCTCCGCTGCTGGAACAGT 510
QY 360 AGAAGATCCACCCCGCTGATTAATAGATCTTTTATATCCATCTTCAAGCTGGGGCC 419
Db 511 GGAAGATCCGACCCCGCTGATTAATAGATCTTTTATATCCATCTTCAAGCTGGGGCC 570
QY 420 AGGAGAAAGTATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 479
Db 571 AGGGAAAGTATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 630
QY 480 GTGTGAGAGTACCAAGAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 539
Db 631 ATGTGAGAGTATCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 687
QY 540 CAAGCCCAATATCATCAGACACGAGACCGGAGTACACACCGATCCACCTTAGCTGA 597
Db 588 TAAGCCCAATATCATCAGACACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 745

```

RESULT 6

```

AAD30156
ID AAD30156 standard; DNA; 2212 BP.
XX
AC AAD30156;
XX
DT 17-MAY-2002 (first entry)
XX
DE Human MCIP associated DNA #2.
XX
KW Muscle calcineurin interacting protein; MCIP; cardiac hypertrophy;
XX heart failure; cardiomyopathy; heart disease; human; gene; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 25..618
FT /tag= a
FT /product= "Human MCIP associated protein #2"

```

W0200204491-A2.

17-JAN-2002.

06-JUL-2001; 2001WO-US21662.

07-JUL-2000; 2000US-216601P.

13-FEB-2001; 2001US-0782953.

(TEXA) UNIV TEXAS SYSTEM.

PA (WILL/) WILLIAMS S R.
XX (ROTH/) ROTHERMEL B.

XX Williams SR, Rothermel B;

XX WPI; 2002-179698/23.

XX P-PSDB; AAE18915.

XX Screening for modulators of muscle calcineurin interacting protein
XX (MCIP) binding, expression or phosphorylation, useful for treating
XX cardiac hypertrophy or heart failure, comprises mixing MCIP,
XX calcineurin and a test compound.

XX Example 1; Page 159-161; 174pp; English.

XX The invention relates to muscle calcineurin interacting proteins (MCIPs)
XX and nucleic acid molecules encoding such proteins. MCIPs form a physical
XX complex with the catalytic subunit of calcineurin and increased levels
XX of MCIPs correspond to a reduced ability of calcineurin to stimulate
XX transcription of certain target genes. The invention also relates to
XX methods for identifying modulators of MCIP binding, expression or
XX phosphorylation. Inhibitors or promoters of MCIP binding to calcineurin
XX may be used for treating cardiac hypertrophy and heart failure.
XX Antibodies to MCIP can be used in characterising the MCIP content of
XX healthy and diseased tissues and subsequently for determining the
XX presence or absence of cardiomyopathy or as predictor of heart disease.
XX The present sequence is human MCIP associated DNA.
XX Note: This sequence has been described as murine MCIP splice variant in
XX the specification, however the sequence seems to be a polynucleotide
XX encoding a MCIP associated protein.

XX Sequence 2212 BP; 588 A; 455 C; 529 G; 640 T; 0 other;

Query Match

Best Local Similarity 68.4%; Score 408.2; DB 24; Length 2212;

Matches 486; Conservative 0; Mismatches 108; Indels 3; Gaps 1;

```

QY 1 ATGGATTTTAGGACTTTAGCTACAAATTTAGCTCCCTGATTGCTGTGGCAACCAT 60
Db 25 ATGGAGGAGTGGAGCTGCGAGGACTGCCAGCGCCACCATCGCTGCACCTGACCGG 84
QY 61 GATGCTTTCAGGCAAGTGAACACAGGCGCAAAATTTGAATCCCTTTCAAGCAATATGAC 120
Db 85 CGCGTGTTCGTGGACGGCTGTGCGCGCCAAATTTGAGTCCCTCTTTAGGACGTATGAC 144
QY 121 AAGGACACCACTTCCAGTATTTTAAAGAGCTTCAAACTGTCCGATGAATTTCAAGCAAC 180
Db 145 AAGGACATCACTTTCAGTATTTTAAAGAGCTTCAAACTGTCCGATGAATTTCAAGCAAC 204
QY 181 CCCTTATCTGAGCGATGCCAGCTGCGGCTGCACAAAGCGAGTTCTTGGGGAAGGAA 240
Db 205 CCCTTCTCCGACGAGATGCCAGCTCCAGCTGCATAGAGTGTGTTCTGGGAAGGAA 264
QY 241 ATGAAGTTTGTATTTTGTCTCAGACTTTTACATAGGAAGTTTCAACCTTGGCTCCGCCAA 300
Db 265 ATGAAGTTTATTTTGTCTCAGACTTTTACATAGGAAGTTTCAACCTTGGCTCCGCCAA 324
QY 301 CCGGACAAACAGTTCTCTCATCTCCCTCCGCTCTCTCCGCTTGGCTGGAACAGTA 360
Db 325 CAGACAAGAGATTTCTGATCTCCCTCCGCTCTCTCCGCTTGGGATGGAACAGTA 384
QY 361 GAAATGCCACCCCGCTCATAAATTTAGCATCTTTTATATGCCATCTCAAGCTGGGGCA 420
Db 385 GAAGATGCGACCCAGTCATAAACTATGATCTCTTATATGCCATCTCAAGCTGGGGCA 444
QY 421 GGAGAGAGTATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
Db 445 GGGGAAAAGTATGAATTTGACAGCGAGCTGACACCACTCCAGCGTGGTGGTCCATGTA 504
QY 481 TGTGAGAGTGAACCAAGAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540
Db 505 TGTGAGAGTGAACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 561

```


CC q22.1-22.2 region of human chromosome 21. An increase in the transient
 CC expression of DSCR1 mRNA in the brains of young rats, compared to
 CC expression levels in the brains of adult rats, suggests an important role
 CC for DSCR1 during the development of the Central Nervous System (CNS), and
 CC that over expression of DSCR1 may be involved in pathogenic abnormalities
 CC of mental retardation and/or heart defects as found in Down's syndrome
 CC patients.
 XX

SQ Sequence 2174 BP; 605 A; 436 C; 502 G; 628 T; 3 other;

Query Match 66.7%; Score 398; DB 20; Length 2174;
 Best Local Similarity 87.6%; Pred. No. 1.4e-113;
 Matches 447; Conservative 0; Mismatches 60; Indels 3; Gaps 1;

QY 88 GCCAAATTTGAATCCCTCTTGAGACATATGACAAAGGACACACCTTCCAGTATTTTAAAG 147

DB 58 GCCAAATTTGAGTCCCTCTTTAGGACGTATGACAAAGGACATCACCTTTCAGTATTTTAAAG 117

QY 148 AGCTTCAAACCGTCCGGATAACTTCAGCAACCCCTTATCTGAGCGGATGCCAGGCTG 207

DB 118 AGCTTCAAACCGTCCGGATAACTTCAGCAACCCCTTCTCCGAGCAGATGCCAGGCTC 177

QY 208 CGGTGCAAGACGAGTTCCTCGGGAAGGAAATGAAGTTGTATTTTGTCTCAGACTTTA 267

DB 178 CAGCTGCATAAGACTGAGTTTCTGGGAAGGAATGAAGTTATATTTTGTCTCAGACTTTA 237

QY 268 CACATAGGAAGTTACACCTGGCTCGGCCCAATCCGCAAAACAGTTTCTCATCTCCCT 327

DB 238 CACATAGGAAGTTACACCTGGCTCGGCCCAATCCGCAAAACAGTTTCTCATCTCCCT 297

QY 328 CCGGCTCTCTCCGTTGGTGGAAACAGTACAGATGCCACCCCGTCCATTAATAC 387

DB 298 CCGGCTCTCGCAGTGGATGGAACAACCTGGAAGATGCGACCCAGTCATAAATAT 357

QY 388 GATCTTTATATGCACTCTCAAGCTGGGCGAGGAGGAATGATGAATGCATGCAGCG 447

DB 358 GATCTTTATATGCACTCTCAAGCTGGGCGAGGAGGAATGATGAATGCATGCAGCG 417

QY 448 ACAGACCCACTCCAGTGGTGGTCCAGTGTGAGAGTGACCAAGAGATGAGGAG 507

DB 418 ACTGACACCACTCCAGCGTGGTGGTCCATGTATGTGAGAGTGATCAAGAGAGGAGGAA 477

QY 508 GAAGAGGAAGAGATGGAGAGAAATGAAGAGACCCCAAGCCCAAAATCATCCAGACAGGGA 567

DB 478 GAAGAGGAA---ATGGAAAGATGAGAGACCTTAAGCCAAAATTTATCCAGACAGGAGG 534

QY 568 CCGAGTACACCGATCCACCTTACCTGTA 597

DB 535 CCGAGTACACCGGATCCACCTCAGCTGA 564

RESULT 9

AAF18328

ID AAF18328 standard; DNA; 2358 BP.

XX

AC AAF18328;

XX

DT 14-MAR-2001 (first entry)

XX

DE Lung cancer associated polynucleotide sequence SEQ ID 347.

XX

KW Human; lung cancer associated protein; neuroprotective; cytoskeletal;
 KW cardioactive; immunomodulatory; muscular active; vulnary;
 KW gastrointestinal; nephrotropic; antiinfective; gynecological;
 KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
 KW proliferative disorder; wound healing; infectious disease; ds.

OS Homo sapiens.

XX

XX WO200055180-A2.

XX

XX 21-SEP-2000.

PD

XX

PF 08-MAR-2000; 2000WO-US05918.

XX

PR 12-MAR-1999; 99US-0124270.

XX

PA (HUMA-) HUMAN GENOME SCI INC.

XX

PA (ROSE/) ROSEN C A.

XX

PI Ruben SM;

XX

DR WPI; 2000-587514/55.

DR P-PSDB; AAB58452.

XX

PT Lung cancer associated gene sequences, referred to as lung cancer

PT antigens, useful for treatment, prevention, and diagnosis of disorders

PT such as lung cancer.

XX

PS Claim 1; Page 805-806; 1425pp; English.

XX

CC Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer

CC associated proteins represented in AAB58106 - AAB58548. Lung cancer

CC associated proteins and polynucleotide sequences, their agonists, and

CC antagonists may have neuroprotective, cytostatic, cardioactive;

CC immunomodulatory; muscular active general; vulnary; gastrointestinal

CC general; nephrotropic; antiinfective; gynecological; or antibacterial

CC activity. The invention also includes antibodies specific for the

CC protein or polynucleotide sequences. The lung cancer associated

CC polynucleotide sequences may be used for detection of lung cancer,

CC chromosome identification, as chromosome markers, and for numerous other

CC diagnostic or research purposes. The proteins may be used to treat

CC disorders such as neural, immune, muscular, reproductive,

CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative

CC disorders. The proteins may also be used in the treatment of wounds and

CC infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and

CC peptide AAB58549 are used in the course of the invention for the

CC identification and characterization of the polynucleotide and protein

XX sequences.

SQ Sequence 2358 BP; 641 A; 482 C; 584 G; 647 T; 4 other;

Query Match 66.2%; Score 395.4; DB 21; Length 2358;

Best Local Similarity 80.9%; Pred. No. 9.4e-113;

Matches 483; Conservative 2; Mismatches 108; Indels 4; Gaps 2;

QY 1 ATGGATTTTAGGGACTTTAGCTTACAAATTTTAGCTCCCTGATTGCTGTGTCGCAACCAT 60

DB 114 ATGGAGGAGGTGGACTCGAGGACCTGCCAGCCACCATCGCTGCACCTGACCCCG 173

QY 61 GATGTCTTCAGCGAAAGTGAGACAGGCCCAAATTTGAATCCCTCTTCAGAACATATGAC 120

DB 174 CGCGTGTTCGTGGACGGCTGTGCGGGCCAAATTTGAGTCCCTCTTTAGGACGATGAC 233

QY 121 AAGGACACCACTTCCAGTATTTTAAGAGCTTCAACGTGTCCGGATAACTTCAGCAAC 180

DB 234 AAGGACATCACCTTTCAGTATTTTAAGAGCTTCAACAGAGTCAGAAATAAATTCAGCAAC 293

QY 181 CCCTTATTCGAGCCGATGCCAGGCTGCCAGTGCACAGACCGAGTTCTCTGGGAAGGAA 240

DB 294 CCCTTCTCGCAGCAGATGCCAGGCTCCAGCTGCATAGACTGAGTTCTCTGGGAAGGAA 353

QY 241 ATGAAGTTGATTTTGTCTCAGCTTTACATAGGAAGTTTACACCTTGGCTCCGCCCAAT 300

DB 354 ATGAAGTTATATTTTGTCTCAGACCTTACATAGGAAGTTTACACCTTGGCTCCGCCCAAT 412

QY 301 CCCGACAAACAGTTTCCCTCATCTCCCTCCGCTCTCTCCGTTGGTGGAAACAAGTA 360

DB 413 CCAGACAAGCAGTTTCTGATCTCCCTCCGCTCTCTCCGAGTGGGATGGAAACAAGTG 472

QY 361 GAAGATGCCACCCCGTCATAAATTAAGATCTTTTATATGATCATCTCAAGCTGGGCCCA 420

DB 473 GAAGATGCCACCCCGTCATAAATTAAGATCTTTTATATGATCTCTCAAGCTGGGCCCA 532

QY 421 GGAGAGAAGTATGAATGTCATGACGACAGACCCCACTCCAGTGTGGTGTCCAGCTG 480

Matches 331; Conservative 0; Mismatches 158; Indels 12; Gaps 1;

QY 21 CTACAAATTTAGTCTCCTGATGCTGTTGGCAAAAGCATGATGTTCTTCAGCGAAAGTGA 80
DB 210 CTGTGATGTTTCCACTCTGTTGCTGTTGGTGGATGTCGAGGTCTTTACCAATCAGGA 269

QY 81 GACCAGGCGCAATTTGAATCCCTCTTCAGACATATGACAGGACACACCTTCCAGTA 140
DB 270 GGTAAAGGAAAAATTTGGGGGACTGTTTCGGACTTATGATGACTGTGTGACGTTCCAGCT 329

QY 141 TTTTAAGAGCTTCAAAACGTTGTCGGGATAAACTTCAGCAACCCCTTATCTGCAGCCGATGC 200
DB 330 ATTAAAGAGTTTCAGAGCGTGTCCGTATAAACTTCAGCAATCTTAATCTGCAGCCGAGC 389

QY 201 CAGGCTCGCGCTGCACAAAGACCGAGTTCTTGGGGAAGGAAATGAAGTGTATTTTCTCA 260
DB 390 TAGGATAGAGCTTTCATCAAAACCCAAATTCAGAGGGGAAAAAATTAAAGCTCTACTTTGCACA 449

QY 261 GACT-----TTACACATAGGAAGTTACACCTGGCTCCGCCCAATCCCGACAA 308
DB 450 GGTTCAGACTCCAGAGACAGATGGACAAACTGCACTTGGCTCCACCCAGCTGCCAA 509

QY 309 ACAGTTCTCTCATCTCCCTCCGCGCTCTCTCCCTGTTGGCTGGAAACAAGTAGAAGATGC 368
DB 510 ACAGTTCTCTCATCTCCGCGCTCTCTCCCTGTTGGCTGGAAACAAGTAGAAGATGC 569

QY 369 CACCCCGCTCAAAATAGCATCTTTATATGCACTTCCAACTCGGGCCAGGAGAGAA 428
DB 570 CAGCCAGTCTCAACTATGACCTCTCTATGCTGTGGCCAAACTAGGACAGGAGAGAA 629

QY 429 GTATGAATGTCATGCGAGCAGACACCCCACTCCAGTGTGGTGTCCAGTGTGTGAGAG 488
DB 630 GTATGAGTCCATGCGAGGAGTGTGCTCACCACCCCAAGTGTGCTGTCAGTGTGCGACAG 689

QY 489 TGACCAAGAGAAATGAGGAGGA 509
DB 690 TGACATAGAGGAAGAGAGGA 710

RESULT 14
ID AAF25338 standard; cDNA; 934 BP.
XX AAF25338;
DT 30-APR-2001 (first entry)
XX Nucleotide sequence of a human detoxification protein.
XX Human; detoxification protein; DETX; cancer; leukaemia; melanoma;
KW adenocarcinoma; autoimmune disorder; inflammatory disorder;
KW rheumatoid arthritis; asthma; atherosclerosis; autoimmune thyroiditis;
KW psoriasis; ulcerative colitis; infection; cell proliferative disorder;
KW actinic keratosis; arteriosclerosis; cirrhosis; hepatitis; ss.
OS Homo sapiens.
XX Key Location/Qualifiers
FH 20..724
CDS /*tag= a
/*product= "detoxification protein"
FT sig_peptide 20..100
FT /*tag= b
XX WO200104305-A2.
XX 18-JAN-2001.
XX 06-JUL-2000; 2000WO-US18509.
XX 07-JUL-1999; 99US-0142678.
XX (INCY-) INCYTE GENOMICS INC.

XX Tang YT, Yue H;
PI WPI; 2001-147193/15.
DR P-PSDB; AAB31788.
XX New human detoxification protein and polynucleotide, useful for
PT diagnosis, prevention and treatment of autoimmune/inflammatory
PT disorders and cell proliferative disorders including cancer -
XX Claim 5; Page 79; 79pp; English.
XX The present sequence encodes a human detoxification polypeptide (DETX).
CC DETX and its (ant)agonists are useful for preventing or treating
CC disorders associated with decreased or increased expression or activity
CC of DETX. DETX polypeptides are useful for screening compounds that
CC specifically binds to DETX and for identifying (ant)agonists.
CC Diseases prevented, treated and diagnosed include cancers (e.g.
CC leukaemia, melanoma, adenocarcinoma, cancers of the brain, bone marrow,
CC breast, kidney, liver, pancreas, prostate and uterus),
CC autoimmune/inflammatory disorders (e.g. rheumatoid arthritis, asthma,
CC atherosclerosis, autoimmune thyroiditis, psoriasis and ulcerative
CC colitis), bacterial, fungal, parasitic infections and cell
CC proliferative disorders (e.g. actinic keratosis, arteriosclerosis,
CC cirrhosis and hepatitis). Anti-DETX antibodies may be used as
CC antagonists, as a targeting or delivery mechanism for bringing
CC pharmaceutical agents into contact with cells or tissues expressing
CC DETX and for diagnosis of DETX-related disorders.
XX Sequence 934 BP; 247 A; 244 C; 206 G; 237 T; 0 other;
SQ Query Match 35.1%; Score 209.6; DB 22; Length 934;
Best Local Similarity 64.8%; Pred. No. 6.1e-55;
Matches 333; Conservative 0; Mismatches 169; Indels 12; Gaps 1;

QY 8 TTAGGAGCTTTAGCTACAAATTTTAGCTCCCTGATGCTGTTGGCAAAAGCATGATGCT 67
DB 138 TCACTGACTTCAATGACTCCCTCCCACTCGTTGTTGGTGCATGTTTCACCACTCAGTGT 197

QY 68 TCAGCGAAAGTGGAGACGAGCGGCAAAATTTGAATCCCTCTTCAGAACATATGACAAGACA 127
DB 198 TTGAAGGAGAGAGACGAGCAAGGAAAAATTTGAGGAGCTGTTTCGGACTTTATGACTGTG 257

QY 128 CCACCTTCCAGTATTTTAAGAGCTTCAAAACGTTCCGAGTAACTTCAGCAACCCCTTAT 187
DB 258 TGAGGTTCCAGCTATTTAAGAGTTTCAGAGCTGTCGCTATAAACTTCAGCAATCCTAAAT 317

QY 188 CTGCAGCCGATGCCAGCTGCGGCTGCCAAGACCGAGTTCCTGGGGAAGGAAATGAAGT 247
DB 318 CTGCAGCCGAGCTAGGATAGAGCTTCATGAACCCCAATTCAGAGGGAAGGAAATTAAGC 377

QY 248 TGTATTTTCTCAGACT-----TTACACATAGGAAGTTTCACACCTCGCTCCGC 295
DB 378 TCTACTTTCCAGGTTTTCAGACTCCAGACAGATGGAGACAACTGCACTTGGCTCCAC 437

QY 296 CCAATCCCGACAAACAGTTCCTCATCTCCCTCCGCGCTCTCTCCCGTGGTGGAAAC 355
DB 438 CCAGGCTGCCAAACAGTTTCTCATCTCGCCCTTCTCTCCCTCCCACTGTTGGCTGCAGC 497

QY 356 AAGTAGAAGATGCCACCCCGCTCATAAATAGATCTTTTATATGTCATCTCCAAGTGG 415
DB 498 CCATCAACGATGCCAGCGAGTCTCAACTATGACCTCTCTATGCTGTGGCCAAACTAG 557

QY 416 GGCCAGGAGAGATGATGAATGATGATGAGCGACAGACCCCACTCCAGTGTGCTGCTCC 475
DB 558 GACCGAGAGAGATGATGAGCTCCATGCGAGGAGCTGAGTCCACCCCAAGTGTGCTGCTGC 617

QY 476 ACCTGTGTGAGAGTGAACCAAGAGAAATGAGGAGGA 509
DB 618 ACCTGTGCGACAGTGCATAGAGGAAGAGAGGA 651

RESULT 15

ID	ABA91457	
AB	ABA91457	standard; cDNA; 3159 BP.
XX		
AC	ABA91457;	
XX		
DT	18-APR-2002	(first entry)
XX		
DE	Human Down syndrome critical region 1-like 1 protein cDNA.	
XX		
KM	Down syndrome critical region 1-like 1; DSCR1L1 alpha; human;	
KW	Down syndrome; Alzheimer's disease; dementia; transgenic;	
KM	neuroprotective; neurotropic; anticonvulsant; diagnosis;	
KM	gene therapy; gene; ss.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	185..952
FT		/*tag= a
FT		/product= "DSCR1L1"
FT		/transl_except= (pos:215..217, aa:Xaa)
FT		/note= "Xaa = unknown"
XX		
PN	W0200204513-A2.	
XX		
XX	17-JAN-2002.	
PD		
XX		
PF	11-JUL-2001; 2001WO-US21982.	
XX		
PR	11-JUL-2000; 2000US-0614474.	
XX		
PA	(INCY-) INCYTE GENOMICS INC.	
XX		
XX	Loring JF, Tingley DW, Edwards CM, Streeter DG;	
P1		
XX		
DR	WPI; 2002-164633/21.	
XX	P-PSDB; AAM50760.	
PT	Novel Down syndrome critical region 1-like 1 protein and nucleic acid	
PT	encoding the protein useful for diagnosis and treatment of Alzheimer's	
XX	disease, Down syndrome and other forms of dementia	
XX		
PS	Claim 2; Page 46-47; 54pp; English.	
XX		
CC	The present sequence is that of cDNA clone incyte ID No: 247500.5	
CC	encoding novel Down syndrome critical region 1-like 1 protein	
CC	(DSCR1L1 alpha, see AAM50760). Northern analysis indicated	
CC	expression of DSCR1L1 alpha in various libraries, with the highest	
CC	abundance in tissues from the nervous system, including tissues	
CC	associated with schizophrenia, Huntington's disease, epilepsy and	
CC	amorphoptic lateral sclerosis. An absence of DSCR1L1 expression	
CC	was observed in 7 of 8 libraries from subjects with Alzheimer's	
CC	disease. A claimed method for detecting differential expression of	
CC	a nucleic acid encoding DSCR1L1 alpha can be used to diagnose Down	
CC	syndrome, Alzheimer's disease and other forms of dementia. A probe	
CC	from such a nucleic acid is useful for identifying naturally	
CC	occurring molecules encoding DSCR1L1 alpha, allelic variants or	
CC	other molecules. The nucleic acids are useful for producing	
CC	transgenic cell lines or organisms which model human disorders.	
CC	They may also be used in gene therapy, and to screen for ligands	
CC	which specifically bind the nucleic acid molecule, such as	
CC	a DNA or RNA molecule, peptide nucleic acid, artificial chromosome,	
CC	peptide, transfection factor, repressor or regulatory molecule,	
CC	for use as a therapeutic.	

Db	366	TCACGACCTTCAAATGACCTCCGCCAACTCGTTGTTGGTCGAATGTTCAACCACTCACTGT	425
Oy	68	TCAGCGAAAGGAGACCAAGGCGCCAAATTTGAATCCCTCTTCAGAACTATGACAAGACA	127
Db	426	TTGAAGGAGAAAGAGACGAGAAATAATTTAGGGACGTGTTTGGACTTATGATGACTGTG	485
Oy	128	CCACCTTCACAGTATTTTAAAGCTTTCAAACGTGTCCGGATPAACTTCAGCAACCCCTTAT	187
Db	486	TGACCTTCCACCTATTATTAAGAGTTTTCAGACGTGTCCGTATPAACTTCAGCAATCTTAAT	545
Oy	188	CTGACGCCGATGCCAGCTGTGGGTGACACAAGACCGAGTTCCTGGGGAAGAAATGAAT	247
Db	546	CTGACGCCCGAGCTPAGATAGAGCTTCATGAAACCCCAATTCAGAGGAAAAAATTAAAG	605
Oy	248	TGATTTTGTCTCAGACT-----TTACATATAGGAAGTTCACACTGACTCCG	295
Db	606	TCTACTTTGGCACAAGTTTCAGACTCCAGACAGACATGAGAGCAAACTGACACTTGCTTCAC	655
Oy	296	CCAAATCCCGACAACAGTTCTCTACTTCCTCCGACTCTCCCTCCGTTGGCTGGAAC	355
Db	666	CCCAAGCTCGCCAAACAGTTTCTCATCTCGCCCCCTTCTCCCACTGTGGCTGGGACG	725
Oy	356	AAGTAGAAGATGCCACCCCGCTCATTAATPACATCTTTATATGCCATCTCCAAAGTGG	415
Db	726	CCATTCACGATGCCACCGCACTCTCAACTATGACTCTCTATGCTGTGGCCAAACTAG	785
Oy	416	GACCGAGAGAGAAATATGAATCTGCATGCAAGACAGACCCCACTCTCCAGTGTGGTGTCC	475
Db	786	GACCGAGAGAAAGTATGAGTCTCATGCAAGGACTGAGTCAACCCCAAGTGTGTGTGTG	845
Oy	476	ACGTGTGTGAGATGACCAAGAAATGAGAGGA	509
Db	846	ACGTGTGTGACATGACATTAAGGAAAGAGGA	879

```
Search completed: December 14, 2002, 15:19:01
Job time : 102.497 secs
```

50 Sequence 3159 BP, 926 A; 657 C; 657 G; 915 T; 4 other;
 Query Match 35.1%; Score 209.6; DB 24; Length 3159;
 Best Local Similarity 64.8%; Pred. No. 1,1e-54;
 Matches 333; Conservative 0; Mismatches 169; Indels 12; Gaps 1

8 TTAGGGACCTTTAGCTACAACTTTTACCTCCCTGATGCTGTGTGGCAAACGATGATGCTCT 67

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2002, 14:11:25 ; Search time 1076.02 Seconds
(without alignments)
16146.861 Million cell updates/sec

Title: US-09-782-953-5

Perfect score: 597

Sequence: 1 atggattttaggacttag.....cacgacccaccttagctga 597

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues .

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.hcg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.ey.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vi.*

30: em.htg_hum.*

31: em.htg_inv.*

32: em.htg_other.*

33: em.htg_mus.*

34: em.htg_pln.*

35: em.htg_rod.*

36: em.htg_mam.*

37: em.htg_vrt.*

38: em.sy.*

39: em.htgo_hum.*

40: em.htgo_mus.*

41: em.htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	597	100.0	597	6	AX365315 Sequence
2	597	100.0	597	10	AF237790 Mus muscu
3	590.6	98.9	2198	10	BC013551 Mus muscu
4	589	98.7	2141	10	AF263329 Mus muscu
5	515.8	86.4	597	6	AX365312 Sequence
6	515.8	86.4	597	10	AF237789 Mus muscu
7	511	85.6	621	10	AF282255 Mus muscu
8	511	85.6	2125	10	AF260717 Mus muscu
9	508.6	85.2	2224	10	AF263240 Mus muscu
10	504.4	84.5	599	6	AX365311 Sequence
11	504	84.4	626	10	AB075973 Rattus no
12	501	83.9	2216	10	CGU60263 Cricetus
13	470.6	78.8	2331	6	AX365321 Sequence
14	470.6	78.8	2346	9	HSU85267 Homo sapien
15	459.6	77.0	2348	6	AX281651 Sequence
16	408.2	68.4	2212	6	AX365324 Sequence
17	408.2	68.4	2227	9	HSU85266 Homo sapien
18	408.2	68.4	2289	9	BC002864 Homo sapien
19	408.2	68.4	2407	9	AC092184 Homo sapien
20	401.8	67.3	562	9	HSU53821 Homo sapien
21	401.6	67.3	2284	9	HSU85265 Homo sapien
22	399.6	66.9	2173	6	AX410694 Sequence
23	399.6	66.9	2173	9	HSU28833 Homo sapien
24	399	66.8	2679	9	AK055845 Homo sapi
25	398	66.7	798	9	AF400429 Homo sapi
26	398	66.7	798	6	AR034241 Sequence
27	214.2	35.9	3184	6	AX329596 Sequence
28	214.2	35.9	3184	6	AX333015 Sequence
29	214.2	35.9	3184	6	AX365327 Sequence
30	214.2	35.9	3184	6	AX420436 Sequence
31	214.2	35.9	3184	9	HUMZAK14 mRNA
32	209.6	35.1	934	6	AX074351 Sequence
33	209.6	35.1	3159	6	AX420425 Sequence
34	209.6	35.1	3261	9	AK090990 Homo sapi
35	206.4	34.6	3240	9	AY034085 Homo sapi
36	206.4	34.6	3253	9	AY034086 Homo sapi
37	201.4	33.7	594	6	AX365318 Sequence
38	201.4	33.7	594	10	AF237791 Mus muscu
39	198.8	33.5	659	10	AF237887 Mus muscu
40	198.2	33.2	3227	10	AB061524 Mus muscu
41	196.8	33.0	3295	10	AB061525 Mus muscu
42	191.2	32.0	720	6	AX365333 Sequence
43	191.2	32.0	720	9	AF176117 Homo sapi
44	191.2	32.0	828	6	AX365330 Sequence
45	191.2	32.0	828	6	AX420437 Sequence

ALIGNMENTS

RESULT 1
AX365315
LOCUS AX365315
DEFINITION Sequence 5 from Patent WO0204491.
ACCESSION AX365315
VERSION AX365315.1
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Williams, S.R. and Rothermel, B.
TITLE Methods and compositions relating to muscle selective calcineurin
interacting protein (mcip)

597 bp DNA linear PAT 15-FEB-2002

JOURNAL

Patent: WO 0204491-A 5 17-JAN-2002;
Bard of Regents, The University of Texas System (US) ; Williams,
Sanders R. (US) ; Rothermel, Beverly (US)
Location/Qualifiers

FEATURES

source

CDS

1..597
/organism="Mus musculus"
/db_xref="taxon:10090"
1..597
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD23809.1"
/db_xref="GI:18697046"
/translation="MDPRDPSYNSFLIACVANDVSESETRAKESLFRDYKDT
FOYFKSRPRVINSNPLSADARLRLHTEFLKEMKLYFAQTLHIGSHLAPND
KQPLISPPASPPVGMKQVEDATPINVDLVAISLKGPEKVELHAATDPTPSVVHV
CESDQENESEEMERMRPKXIOTRRPPTPHLS"

BASE COUNT 170 a 156 c 142 g 129 t
ORIGIN

Query Match 100.0%; Score 597; DB 6; Length 597;

Best Local Similarity 100.0%; Pred. No. 5.1e-162; Mismatches 0; Indels 0; Gaps 0;

Matches 597; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CDS

1..597
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="16"
1..597
/note="similar to the EST with GenBank Accession Number
AA200984; similar to Homo sapiens DSCR1"
/product="myocyte-enriched calcineurin interactin protein
1 splice variant 4"
/protein_id="AA63486.1"
/db_xref="GI:7542529"
/translation="MDPRDPSYNSFLIACVANDVSESETRAKESLFRDYKDT
FOYFKSRPRVINSNPLSADARLRLHTEFLKEMKLYFAQTLHIGSHLAPND
KQPLISPPASPPVGMKQVEDATPINVDLVAISLKGPEKVELHAATDPTPSVVHV
CESDQENESEEMERMRPKXIOTRRPPTPHLS"

BASE COUNT 170 a 156 c 142 g 129 t
ORIGIN

Query Match 100.0%; Score 597; DB 10; Length 597;

Best Local Similarity 100.0%; Pred. No. 5.1e-162; Mismatches 0; Indels 0; Gaps 0;

Matches 597; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORGANISM

Mus musculus

REFERENCE

AUTHORS

TITLE

A Protein Encoded within the Down Syndrome Critical Region is
Enriched in Striated Muscles and Inhibits Calcineurin Signaling

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

```

Db 481 TGTGAGGTGACCAAGAGAAATGAGGAGGAGGAGAGATGAGAGAAATGAAGAGACCC 540
QY 541 AAGCCCAAAATCATCCAGACACGAGACCGGAGTACACACCGATCCACTTAGCTGA 597
Db 541 AAGCCCAAAATCATCCAGACACGAGACCGGAGTACACACCGATCCACTTAGCTGA 597

RESULT 3
BC013551 2198 bp mRNA linear ROD 07-AUG-2002
LOCUS Mus musculus, Down syndrome critical region homolog 1 (human),
DEFINITION clone MGC:19348 IMAGE:4236038, mRNA, complete cds.
ACCESSION BC013551
VERSION BC013551.1 GI:15488840
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
Strausberg, R.
Direct Submission
Submitted (04-SEP-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www.shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 25 Row: m Column: 6.

FEATURES
source
1..2198
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="FVB/N"
/clone="MGC:19348 IMAGE:4236038"
/tissue_type="Kidney, normal. 5 month old male mouse."
/clone_lib="NCI CGAP_Kid14"
/lab_hosts="DH10B"
/notes="Vector: pCMV-SPORT6"
77..673
/codon_start=1
/product="Down syndrome critical region homolog 1 (human)"
/protein_id="AAH13551.1"
/db_xref="GI:15488841"
/db_xref="LocusID:54720"
/translation="MHFRDFSYNFSLLIACVANDVDFSESTRAKFESLFRTYDKDFTT
KQYKFSKFRVFNFSPLSAADARLRLHKTEFLGKEMKLYPAQTLHGSHLAPNPD
KQFLISPPASPPGVKQVEDATPVINDYLLVLSKLGPKGYELHNAITDTPSVVVHV
CESDQNEEEEMEMKPKKIOTRPEYPIHLS"

BASE COUNT 533 a 539 c 569 g 557 t
ORIGIN

Query Match 98.9%; Score 590.6; DB 10; Length 2198;
Best Local Similarity 99.3%; Pred. No. 4.2e-160;
Matches 593; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATCGATTATAGGACCTTAGCTACAATTTAGCTCCCTGATCTGTGTGGCAACGAT 60
Db 77 ATGCATTTAGGACCTTAGCTACAATTTAGCTCCCTGATCTGTGTGGCAACGAT 136

```

```

QY 61 GATGTCCTTACGCGAAAGTGAAGACGAGGCCAAATTTGAATCCCTCTTTCAGAACATATGAC 120
Db 137 GATGTCCTTACGCGAAAGTGAAGACGAGGCCAAATTTGAATCCCTCTTTCAGAACATATGAC 196
QY 121 AAGGACACCAACCTTCCAGTATTTTAAAGAGCTTCAAAAGTGTCCGGATAAATTTTCAGCAAC 180
Db 197 AAGGACACCAACCTTCCAGTATTTTAAAGAGCTTCAAAAGTGTCCGGATAAATTTTCAGCAAC 256
QY 181 CCCTTATCTGACGCGGATGCCAGGCTGCGGCTGCACAAGACCGAGTTCCCTGGGGAAGGAA 240
Db 257 CCCTTATCTGACGCGGATGCCAGGCTGCGGCTGCACAAGACCGAGTTCCCTGGGGAAGGAA 316
QY 241 ATGAAGTTGTATTTGCTCAGACTTTACACATAGGAAGTTTACACCTGGTCCGCCCAAT 300
Db 317 ATGAAGTTGTATTTGCTCAGACTTTACACATAGGAAGTTTACACCTGGTCCGCCCAAT 376
QY 301 CCGGACAAACAGTTCCTCATCTCCCTCCGCGCTCTCTCCCTGGCTGGGAAACAAGTA 360
Db 377 CCGGACAAACAGTTCCTCATCTCCCTCCGCGCTCTCTCCCTGGCTGGGAAACAAGTA 436
QY 361 GAAGATGCCACCCCGCTCATTAATTAAGATCTTTTATATGCCATCTCCAAGCTGGGGCCA 420
Db 437 GAAGATGCCACCCCGCTCATTAATTAAGATCTTTTATATGCCATCTCCAAGCTGGGGCCA 496
QY 421 GGAGAGAAGTATGAAGCTGCATGCGAGCAGACGCCACCTCCAGTGTGTGTTCCACGTG 480
Db 497 GGAGAGAAGTATGAAGCTGCATGCGAGCAGACGCCACCTCCAGTGTGTGTTCCACGTG 556
QY 481 TGTGAGAGTGACCAAGAGAAATGAGGAGGAGGAGAGATGAGAGAAATGAAGAGACCC 540
Db 557 TGTGAGAGTGACCAAGAGAAATGAGGAGGAGGAGAGATGAGAGAAATGAAGAGACCC 616
QY 541 AAGCCCAAAATCATCCAGACACGAGACCGGAGTACACACCGATCCACTTAGCTGA 597
Db 617 AAGCCCAAAATCATCCAGACACGAGACCGGAGTACACACCGATCCACTTAGCTGA 673

RESULT 4
AF263239 2141 bp mRNA linear ROD 30-MAY-2000
LOCUS Mus musculus calcineurin inhibitor mRNA, complete cds,
DEFINITION alternatively spliced.
ACCESSION AF263239
VERSION AF263239.1 GI:8102011
KEYWORDS Mus musculus.
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
Fuentes, J.J., Pritchard, M.A., Pucharcos, C. and Estivill, X.
Down syndrome candidate region 1 (Dscr1), one of three
alternatively spliced exon 1 transcripts
Unpublished
2 (bases 1 to 2141)
Fuentes, J.J., Pritchard, M.A., Pucharcos, C. and Estivill, X.
Direct Submission
Submitted (02-MAY-2000) Centre de Genetica Medica i Molecular, IRO,
Avda. Castelldefels Km. 2.7, L'Hospitalet de Llobregat, Barcelona
08907, Spain
Location/Qualifiers
1..2141
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="16"
31..627
/notes="Down syndrome candidate region 1; Dscr1;
alternatively spliced exon 1"
/codon_start=1
/product="calcineurin inhibitor"
/protein_id="AAF72701.1"
/db_xref="GI:8102012"

FEATURES
source
1..2141
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="16"
31..627
/notes="Down syndrome candidate region 1; Dscr1;
alternatively spliced exon 1"
/codon_start=1
/product="calcineurin inhibitor"
/protein_id="AAF72701.1"
/db_xref="GI:8102012"

```

/translation="MHPRDYSYNPSSLIACVANDVFESEETRAKESLFRDYKDT
FOYFKSPKRVINFSNPLSADARLKHTEFLGKEMKLYAQTLHSSHLAPND
KQFLSPSPSPVGMKQVEDATPVINIDLVAISKLGPEKYEHLHAADPTPSVVHV
CESDENEEREEEMERKRPKXITOTRREPTPIHLS"

BASE COUNT 505 a 527 c 558 g 548 t 3 others
ORIGIN

Query Match 98.7%; Score 589; DB 10; Length 2141;
Best Local Similarity 99.2%; Pred. No. 1.2e-159;
Matches 592; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 ATGATTTTGGGAGCTTACTACATTTTACCTCCCTGATTGCTGTGTGGCAACCAT 60
DB 31 ATGATTTTGGGAGCTTACTACATTTTACCTCCCTGATTGCTGTGTGGCAACCAT 90
OY 61 GATGCTTTCAGCGAAAGTGAGACGAGGCCAAATTGAAATCCCTCTTCAGAAATATGAC 120
DB 91 GATGCTTTCAGCGAAAGTGAGACGAGGCCAAATTGAAATCCCTCTTCAGAAATATGAC 150
OY 121 AAGGACACCACTTCAGATTTTAAAGAGCTTCAACGTGTCGGATTAACCTTCAGCAAC 180
DB 151 AAGGACACCACTTCAGATTTTAAAGAGCTTCAACGTGTCGGATTAACCTTCAGCAAC 210
OY 181 CCTTATCTGACAGCGATCCAGGCTGCGCTGCAACAGACCGAGTCTGTGGGAAGAA 240
DB 211 CCTTATCTGACAGCGATCCAGGCTGCGCTGCAACAGACCGAGTCTGTGGGAAGAA 270
OY 241 ATGAACTGTATTTTCTCAGACTTACATATGAAAGTTTCAACCTGCTCCGGCCCAT 300
DB 271 ATGAACTGTATTTTCTCAGACTTACATATGAAAGTTTCAACCTGCTCCGGCCCAT 330
OY 301 CCGGACAAACAGTCTCTCCCTCCGGCTCTCTCCCTGCTGCGGTGGAACAAGTA 360
DB 331 CCGGACAAACAGTCTCTCCCTCCGGCTCTCTCCCTGCTGCGGTGGAACAAGTA 390
OY 361 GAAGATGCCACCCCGTCATTAATTCAGATCTTTTATATGCCATCTTCAAGCTGGGGCCA 420
DB 391 GAAGATGCCACCCCGTCATTAATTCAGATCTTTTATATGCCATCTTCAAGCTGGGGCCA 450
OY 421 GGAAGAAGTATGAATGTCATGAGGAGACAGCCCACTCCAGTGGTGGTCCAGTG 480
DB 451 GGAAGAAGTATGAATGTCATGAGGAGACAGCCCACTCCAGTGGTGGTCCAGTG 510
OY 481 TGTGAGAGTACCAAGAGATGAGAGAAAGAGAAAGATGGAAGAAAGAGAGCC 540
DB 511 TGTGAGAGTACCAAGAGATGAGAGAAAGAGAAAGATGGAAGAAAGAGAGCC 570
OY 541 AAGCCCAAAATCATCCAGACCGAGACCGAGATCACACCCATCCACCTCAGCTGA 597
DB 571 AAGCCCAAAATCATCCAGACCGAGACCGAGATCACACCCATCCACCTCAGCTGA 627

RESULT 5
AX365312 597 bp DNA linear PAT 15-FEB-2002
LOCUS AX365312
DEFINITION Sequence 2 from Patent WO0204491.
ACCESSION AX365312
VERSION AX365312.1 GI:18697043
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 Williams, S.R. and Rothenmel, B.
AUTHORS Methods and compositions relating to muscle selective calcineurin
TITLE interacting protein (mciip)
JOURNAL Patent: WO 0204491-A 2 17-JAN-2002;
BOARD OF Regents, The University of Texas System (US); Williams,
SANDERS R. (US); Rothenmel, Beverly (US)
FEATURES
SOURCE Location/Qualifiers
1..597
/organism="Mus musculus"

CDS /db_xref="caxon:10090"
1..597
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD23808.1"
/db_xref="GI:18697044"

BASE COUNT 160 a 170 c 151 g 116 t
ORIGIN

Query Match 86.4%; Score 515.8; DB 6; Length 597;
Best Local Similarity 94.4%; Pred. No. 1.8e-138;
Matches 535; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

OY 31 AGCTCCCTGATTGCTTGTGTGCAACGATGATGCTTTCAGCAAGAGTGACAGGAGCC 90
DB 31 AGCTCCCTGATTGCTTGTGTGCAACGATGATGCTTTCAGCAAGAGTGACAGGAGCC 90
OY 91 AAATTTGAATCCCTCTTCAGAAATATGACAGACACCACTTCAGATATTTTAAAGAC 150
DB 91 AAATTTGAATCCCTCTTCAGAAATATGACAGACACCACTTCAGATATTTTAAAGAC 150
OY 151 TTCAAGCTGTCCGGAATAACTTCAGCAACCCCTTATCTGCAAGCCGATGCGAGCTGCGG 210
DB 151 TTCAAGCTGTCCGGAATAACTTCAGCAACCCCTTATCTGCAAGCCGATGCGAGCTGCGG 210
OY 211 CTGCAAGAGCCGAGTCTCGGGGAAGAAATGAAAGTTGATTTTGTCTCAGACTTTAC 270
DB 211 CTGCAAGAGCCGAGTCTCGGGGAAGAAATGAAAGTTGATTTTGTCTCAGACTTTAC 270
OY 271 ATGAAAGTTTCAACCTGCTCCGCCCAATCCGACAAAGTTCTCATCTCCCTCCG 330
DB 271 ATGAAAGTTTCAACCTGCTCCGCCCAATCCGACAAAGTTCTCATCTCCCTCCG 330
OY 331 GCCTCTCTCCCTGCTGAGAAACAAGTAAATGATCCACCCCGTCATTAATTAAGAT 390
DB 331 GCCTCTCTCCCTGCTGAGAAACAAGTAAATGATCCACCCCGTCATTAATTAAGAT 390
OY 391 CTTTATATGTCATCTCCAGCTGGGGCCAGAGAGATGATGAACTGCACTGACGACA 450
DB 391 CTTTATATGTCATCTCCAGCTGGGGCCAGAGAGATGATGAACTGCACTGACGACA 450
OY 451 GAACCCACTCCCAAGTGTGTGTCAGCTGTCAGAGTGCACCAAGAGATGAGAGAA 510
DB 451 GAACCCACTCCCAAGTGTGTGTCAGCTGTCAGAGTGCACCAAGAGATGAGAGAA 510
OY 511 GAGGAAGAGATGAGAGAAATGAAGAGACCAAGCCCAAAATCATTCAGACAGAGAGCC 570
DB 511 GAGGAAGAGATGAGAGAAATGAAGAGACCAAGCCCAAAATCATTCAGACAGAGAGCC 570
OY 571 GAGTACACACCGATCCACCTTAAGCTGA 597
DB 571 GAGTACACACCGATCCACCTTAAGCTGA 597

RESULT 6
AF237789 597 bp mRNA linear ROD 12-APR-2000
LOCUS AF237789
DEFINITION Mus musculus myocyte-enriched calcineurin interactin protein 1
ACCESSION AF237789
VERSION AF237789.1 GI:7542525
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 Rothenmel, B.A., Vega, R.B., Yang, J., Wu, H., Baaseel-Duby, R.S. and
AUTHORS Williams, R.S.

TITLE * A Protein Encoded within the Down Syndrome Critical Region is Enriched in Striated Muscles and Inhibits Calcineurin Signaling
J. Biol. Chem. (2000) In press
REFERENCE 2 (bases 1 to 597)
AUTHORS Rothermel, B.A., Vega, R.B., Yang, J., Wu, H., Bassel-Duby, R.S. and Williams, R.S.
TITLE Direct Submission
JOURNAL Submitted (23-PEB-2000) Internal Medicine, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, TX 75390, USA

FEATURES
source
1..597
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="16"
1..597
/note="similar to Homo sapiens DSCR1; MC1P1; similar to EST sequence of GenBank Accession Number AA734360"
/codon_start=1
/product="myocyte-enriched calcineurin interactin protein 1 splice variant 1"
/protein_id="AAF63485.1"
/db_xref="GI:7542526"
/translation="MEEVDLQDLPSATIACHLDPRVFDGLCRAKFESLFTYDKDTP
QYFKSFRVRINFNSPLSADARLHKTEFLGKEMKLYFAOTLHGSSHLAPPND
KQFLISPPSPVGMKQVEDATPVINDLLYAIKLGPGKELHAAATDTPSVVVHV
CESQNEEBEERMERKPKKIOTRRPEYTPHLS"
BASE COUNT 160 a 170 c 151 g 116 t
ORIGIN

Query Match 86.4%; Score 515.8; DB 10; Length 597;
Best Local Similarity 94.4%; Pred. No. 1.8e-138;
Matches 535; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
QY 31 AGCTCCCTGATTGCTGTGGCAACGATGATGCTTTCAGCGAAGTGACACGAGGCC 90
DB 31 AGCGCCACCATCGCTGCACCTGACCGCGCGTGTTCGTGGAGCGCTGTGCGGCGCC 90
QY 91 AAATTGTAATCCCTCTTCAGAACATATGACAGGACGACACCTTCCAGTATTTAAAGC 150
DB 91 AAATTGTAATCCCTCTTCAGAACATATGACAGGACGACACCTTCCAGTATTTAAAGC 150
QY 151 TTCAAACTGTCCGGATAACTTCAGCAACCCCTTATCTGAGCGGATGCCAGCTCGG 210
DB 151 TTCAAACTGTCCGGATAACTTCAGCAACCCCTTATCTGAGCGGATGCCAGCTCGG 210
QY 211 CTGCACAGACCGAGTCTCTGGGAAGAAATGAAGTTGATTTTGTCTCAGACTTTACAC 270
DB 211 CTGCACAGACCGAGTCTCTGGGAAGAAATGAAGTTGATTTTGTCTCAGACTTTACAC 270
QY 271 ATAGGAAGTTCCACACTGGCTCGGCCCAATCCGACAAACAGTTCCTCATCTCCCTCG 330
DB 271 ATAGGAAGTTCCACACTGGCTCGGCCCAATCCGACAAACAGTTCCTCATCTCCCTCG 330
QY 331 GCCTCTCTCCGTTGGCTGGAAACAGTAGAAGATGCCACCCCGCTCATAAATACGAT 390
DB 331 GCCTCTCTCCGTTGGCTGGAAACAGTAGAAGATGCCACCCCGCTCATAAATACGAT 390
QY 391 CTTTATATGCCATCTCAAGCTGGGCGCAGGAGAGATGATGAATGATGATGATGATGATG 450
DB 391 CTTTATATGCCATCTCAAGCTGGGCGCAGGAGAGATGATGAATGATGATGATGATGATG 450
QY 451 GACCCCACTCCCAAGT 510
DB 451 GACCCCACTCCCAAGT 510
QY 511 GAGGAAGATGAGAGAAATGAAGAGACCCCAAGCCCAAAATCATCCAGACCGAGACCG 570
DB 511 GAGGAAGATGAGAGAAATGAAGAGACCCCAAGCCCAAAATCATCCAGACCGAGACCG 570
QY 571 GAGTACACCGATCCACCTTACCTGA 597
DB 571 GAGTACACCGATCCACCTTACCTGA 597

RESULT 7
AF282255
LOCUS
DEFINITION
complete cds.
ACCESSION
AF282255
VERSION
AF282255.1
KEYWORDS
GI:9652249
SOURCE
Mus musculus.
ORGANISM
Mus musculus

REFERENCE
1 (bases 1 to 621)
AUTHORS
Strippoli, P., Petrini, M., Lenzi, L., Carinci, P. and Zannotti, M.
TITLE
The murine DSCR1-like (Down syndrome candidate region 1) gene family: conserved syntenic with the human orthologous genes
JOURNAL
MEDLINE
20534792
PUBMED
11080588
REFERENCE
2 (bases 1 to 621)
AUTHORS
Strippoli, P., Petrini, M., Lenzi, L., Carinci, P. and Zannotti, M.
TITLE
Direct Submission
JOURNAL
Submitted (24-JUN-2000) Istituto di Istologia ed Embriologia Generale, Universita di Bologna, Via Belmeloro, 8, Bologna, BO 40126, Italy

FEATURES
Location/Qualifiers
1..621
/organism="Mus musculus"
/strain="BALB/c"
/db_xref="taxon:10090"
/chromosome="16"
/sex="male"
/tissue_type="brain"
1..621
/gene="Dscr1"
8..604
/gene="Dscr1"
/note="similar to the Homo sapiens Down syndrome candidate region 1 protein"
/codon_start=1
/product="Down syndrome candidate region 1 protein"
/protein_id="AAF91461.1"
/db_xref="GI:9652250"
/translation="MEEVDLQDLPSATIACHLDPRVFDGLCRAKFESLFTYDKDTP
QYFKSFRVRINFNSPLSADARLHKTEFLGKEMKLYFAOTLHGSSHLAPPND
KQFLISPPSPVGMKQVEDATPVINDLLYAIKLGPGKELHAAATDTPSVVVHV
CESQNEEBEERMERKPKKIOTRRPEYTPHLS"
BASE COUNT 166 a 179 c 159 g 117 t
ORIGIN

Query Match 85.6%; Score 511; DB 10; Length 621;
Best Local Similarity 93.8%; Pred. No. 4.5e-137;
Matches 532; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
QY 31 AGCTCCCTGATTGCTGTGGCAACGATGATGCTTTCAGCGAAGTGACACGAGGCC 90
DB 31 AGCGCCACCATCGCTGCACCTGACCGCGCGTGTTCGTGGAGCGCTGTGCGGCGCC 97
QY 91 AAATTGTAATCCCTCTTCAGAACATATGACAGGACGACACCTTCCAGTATTTAAAGC 150
DB 91 AAATTGTAATCCCTCTTCAGAACATATGACAGGACGACACCTTCCAGTATTTAAAGC 157
QY 151 TTCAAACTGTCCGGATAACTTCAGCAACCCCTTATCTGAGCGGATGCCAGCTCGG 210
DB 151 TTCAAACTGTCCGGATAACTTCAGCAACCCCTTATCTGAGCGGATGCCAGCTCGG 217
QY 211 CTGCACAGACCGAGTCTCTGGGAAGAAATGAAGTTGATTTTGTCTCAGACTTTACAC 270
DB 211 CTGCACAGACCGAGTCTCTGGGAAGAAATGAAGTTGATTTTGTCTCAGACTTTACAC 277
QY 271 ATAGGAAGTTCCACACTGGCTCGGCCCAATCCGACAAACAGTTCCTCATCTCCCTCG 330

```

Db 278 ATAGAAATTACACCTGCTCCGCCCAATCCCAACAAACAGTTCCTCATCTCCCTCCG 337
Oy 331 GCCTCTCCCTCCGTTGGCTGGAAACAGTAGAAGATGCGACCCCGCTCATTAATTAGCAT 390
Db 338 GCCTCTCTCCGTTGGCTGGAAACAGTAGAAGATGCGACCCCGCTCATTAATTAGCAT 397
Oy 391 CTTTATATGCGATCTCCAGCTGGGGCCAGAGAGATGAACTGATGACAGCGACA 450
Db 398 CTTTATATGCGATCTCCAGCTGGGGCCAGAGAGATGAACTGATGACAGCGACA 457
Oy 451 GACCCCACTCCAGTGTGGTGTCCAGCTGTGTGAGATGACCAAGAGATGAGAGAA 510
Db 458 GACACCACTCCAGTGTGGTGTCCAGCTGTGTGAGATGACCAAGAGATGAGAGAA 517
Oy 511 GAGGAAGAGATGAGAGAAATGAAGAGACCCCAAAATATATCCAGACAGGAGACG 570
Db 518 GAGGAAGAGATGAGAGAAATGAAGAGACCCCAAAATATATCCAGACAGGAGACG 577
Oy 571 GAGTACACACCGATCCACTTAAGCTGA 597
Db 578 GAGTACACACCGATCCACTTAAGCTGA 604

RESULT 8
AF260717 2125 bp mRNA linear ROD 22-MAR-2001
LOCUS Mus musculus Down syndrome candidate region 1 (Dscr1) mRNA, complete
DEFINITION
cde.
ACCESSION AF260717
VERSION AF260717.1 GI:7839596
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Metazoa; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2125)
AUTHORS Casas C., Martinez S., Pritchard M.A., Fuentes J.J., Nadal M.,
Guinera J., Arbones M., Florez J., Soriano E., Estivill X. and
Alcantara S.
Dscr1, a novel endogenous inhibitor of calcineurin signaling, is
expressed in the primitive ventricle of the heart and during
neurogenesis
Mech. Dev. 101 (1-2), 289-292 (2001)

JOURNAL MEDLINE 2152920
PUBMED 11231093
REFERENCE 2 (bases 1 to 2125)
AUTHORS Fuentes J.J., Pritchard M., Pucharcos C. and Estivill X.
TITLE Down syndrome candidate region 1 (Dscr1), one of three
alternatively spliced exon 1
unpublished
JOURNAL 3 (bases 1 to 2125)
AUTHORS Fuentes J.J., Pritchard M., Pucharcos C. and Estivill X.
TITLE Direct Submission
Submitted (26-APR-2000) Centre de Genetica Medica i Molecular, IRO,
Avia, Castelldefels Km. 2,7, L'Hospitalet de Llobregat, Barcelona
08907, Spain

FEATURES
Source
1..2125
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="16"
1..2125
/gene="Dscr1"
15..611
/gene="Dscr1"
/note="Calcineurin inhibitor"
/codon_start=1
/product="Down syndrome candidate region 1"
/protein_id="AA070343.1"
/db_xref="GI:7839597"
/translation="MSEVDIADPSATLACHLDPRVVDGLGAKFESLFRYDQDTT
FOYKFSKRVIRINFSNPLSADARLRLHTEFLGKEMKLYFAQTLLIGSSHLAPRPD

```

```

BASE COUNT 487 a 536 c 563 g 536 t
ORIGIN
Query Match 85.6%; Score 511; DB 10; Length 2125;
Best Local Similarity 93.8%; Pred. No. 5,2e-137;
Matches 532; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

Oy 31 AGCTCCGATGCTGCTGTGGCAACGATGATGTCCTTCAGCGAAAGTGAACAGGCGC 90
Db 45 AGCGCCACCATGCCCTGCGACCTGGACCGCGGGTGTGTGAGAGCGGCTGCGGGCC 104
Oy 91 AAATTGAAATCCCTCTTCAGAACATATGACAAAGACACCACTTCAGATTTTAAAGAC 150
Db 105 AAATTGAAATCCCTCTTCAGAACATATGACAAAGACACCACTTCAGATTTTAAAGAC 164
Oy 151 TTCAAAGTGTCCGGATTAACCTTCAGCAACCCCTTATCTGACGCGATGCGAGGCTGGG 210
Db 165 TTCAAAGTGTCCGGATTAACCTTCAGCAACCCCTTATCTGACGCGATGCGAGGCTGGG 224
Oy 211 CTGCAACAAGCCGAGTTCCTGGGGAAGAAATGAAAGTTGATTTTGTCTGACACTTTAC 270
Db 225 CTGCAACAAGCCGAGTTCCTGGGGAAGAAATGAAAGTTGATTTTGTCTGACACTTTAC 284
Oy 271 ATAGAAATTACACCTGCTCCGCCCAATCCGACAAAGTTCCTCATCTCCCTCCG 330
Db 285 ATAGAAATTACACCTGCTCCGCCCAATCCGACAAAGTTCCTCATCTCCCTCCG 344
Oy 331 GCCTCTCTCCGTTGGCTGGAAACAGTAGAAGATGCGACCCCGCTCATTAATTAGCAT 390
Db 345 GCCTCTCTCCGTTGGCTGGAAACAGTAGAAGATGCGACCCCGCTCATTAATTAGCAT 404
Oy 391 CTTTATATGCGATCTCCAGCTGGGGCCAGAGAGATGAACTGATGACAGCGACA 450
Db 405 CTTTATATGCGATCTCCAGCTGGGGCCAGAGAGATGAACTGATGACAGCGACA 464
Oy 451 GACCCCACTCCAGTGTGGTGTCCAGCTGTGTGAGATGACCAAGAGATGAGAGAA 510
Db 465 GACACCACTCCAGTGTGGTGTCCAGCTGTGTGAGATGACCAAGAGATGAGAGAA 524
Oy 511 GAGGAAGAGATGAGAGAAATGAAGAGACCCCAAAATATATCCAGACAGGAGACG 570
Db 525 GAGGAAGAGATGAGAGAAATGAAGAGACCCCAAAATATATCCAGACAGGAGACG 584
Oy 571 GAGTACACACCGATCCACTTAAGCTGA 597
Db 585 GAGTACACACCGATCCACTTAAGCTGA 611

RESULT 9
AF263240 2224 bp mRNA linear ROD 30-MAY-2000
LOCUS Mus musculus calcineurin inhibitor mRNA, complete cds,
DEFINITION
alternatively spliced.
ACCESSION AF263240
VERSION AF263240.1 GI:8102013
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Metazoa; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2224)
AUTHORS Fuentes J.J., Pritchard M.A., Pucharcos C. and Estivill X.
TITLE Down syndrome candidate region 1 (Dscr1), one of three
alternatively spliced exon 1 transcripts
unpublished
JOURNAL 2 (bases 1 to 2224)
AUTHORS Fuentes J.J., Pritchard M.A., Pucharcos C. and Estivill X.
TITLE Direct Submission
Submitted (02-MAY-2000) Centre de Genetica Medica i Molecular, IRO,
Avia, Castelldefels Km. 2,7, L'Hospitalet de Llobregat, Barcelona
08907, Spain

```


Db	384	ATGAAGTTTATTATTTTGCTCAGACCTTTACATAGGAAGCTCACACTGGCTCCGCCAAAT	443
Qy	301	CCGCACAAAAGTTTCCTCATCTCCCTCCGCGCTCTCTCTCCCTTGGCTGGAAACAAGTA	360
Db	444	CCAGACAAGCAAGTTTCGTATCTCCCTCCGCGCTCTCCGCGCTTCCGCGAGTGGATGGAACAAGTG	503
Qy	361	GAGATGCCACCCCCTGCATAAATACGATCTTTTATATGCCATCTCCAAGCTGGGGCCA	420
Db	504	GAGATGCCACCCCAGTCATAAACATATGATCTCTTATATGCCATCTCCAAGCTGGGGCCA	563
Qy	421	GGAGAGAAGTAGTAACATGCATGCAGCGACAGCCCCACCTCCAGTGTGTGTGTCACAGTG	480
Db	564	GGGAAAAGTAGTAATTCACAGCGAGCGACTGACACCACCTCCAGCTGTGTGTGTCATGTA	623
Qy	481	TGTGAGAGTGACCAAGAAGATGAGGAGGAGGAAGAGATGGAGAGATGAGAGACCC	540
Db	624	TGTGAGAGTGATCAAGAGAAGGAGGAGAGAGAA--ATGAAAGAATGAGGAGACCT	680
Qy	541	AAGCCCAAATCATCCAGACACCGAGACCGGAGTACACACCGATCCACCTTAGCTGA	597
Db	681	AAGCAAAAAATTATCCAGACGAGGAGCGCGAGTACACGCGATCCACCTCAGCTGA	737
RESULT	14		
HSU85267			
LOCUS	HSU85267	2346 bp	mRNA linear PRI 19-APR
DEFINITION	Homo sapiens down syndrome candidate region 1 (DSCR1) gene, alternative exon 1, complete cds.		
ACCESSION	U85267		
VERSION	U85267.2 GI:7596915		
KEYWORDS			
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	Fuentes,J.J., Pritchard,M.A., Planas,A.M., Bosch,A., Ferrer,I. Estivill,X.		
JOURNAL	A new human gene from the Down syndrome critical region encodes proline-rich protein highly expressed in fetal brain and heart		
MEDLINE	Hum. Mol. Genet. 4 (10), 1935-1944 (1995)		
PUBMED	96121593		
REFERENCE	8595418		
AUTHORS	2 (bases 1 to 2346)		
TITLE	Fuentes,J.J., Pritchard,M.A. and Estivill,X. Genomic organization, alternative splicing, and expression pattern of the DSCR1 (Down syndrome candidate region 1) gene		
JOURNAL	Genomics 44 (3), 358-361 (1997)		
MEDLINE	97468152		
PUBMED	9325060		
REFERENCE	3 (bases 1 to 2346)		
AUTHORS	Fuentes,J.J., Pritchard,M. and Estivill,X.		
TITLE	Direct Submision		
JOURNAL	Submitted (14-JAN-1997) Genetics Molecular, Institut de Recerca Oncologica, Autovia de Castelldefels, Km 2.7, Hospitalet, Barce		
REFERENCE	4 (bases 1 to 2346)		
AUTHORS	Fuentes,J.J., Pritchard,M. and Estivill,X.		
TITLE	Direct Submision		
JOURNAL	Submitted (19-APR-2000) Genetics Molecular, Institut de Recerca Oncologica, Autovia de Castelldefels, Km 2.7, Hospitalet, Barce		
REMARK	Sequence update by submitter		
COMMENT	On Apr 19, 2000 this sequence version replaced gi:3612867.		
FEATURES	Location/Qualifiers		
source	1..2346 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="21" /map="21q22.1-q22.2"		
gene	1..2346 /gene="DSCR1"		
CDS	14..737		

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2002, 14:11:25 ; Search time 4201.35 Seconds
(without alignments)
16146.861 Million cell updates/sec

Title: US-09-782-953-11
Perfect score: 2331
Sequence: 1 tttttttttccccggaggt.....gggcacagtatgctcttacc 2331

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_hg_hum.*

31: em_hg_inv.*

32: em_hg_other.*

33: em_hg_mus.*

34: em_hg_pln.*

35: em_hg_rod.*

36: em_hg_nam.*

37: em_hg_vrt.*

38: em_sy.*

39: em_hgo_hum.*

40: em_hgo_mus.*

41: em_hgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	2331	100.0	2331	6	AX365321
2	2331	100.0	2346	9	HSU85267
3	2318.4	99.5	2348	6	AX281651
4	2109	90.5	2212	6	AX365324
5	2109	90.5	2227	9	HSU85266
6	2107.4	90.4	2289	9	BC002864
7	2105.8	90.3	2407	9	AK092184
8	2104	90.3	2284	9	HSU85265
9	2101	90.1	2173	6	AX410694
10	2101	90.1	2173	9	HSU28833
11	2093.6	89.8	2679	9	AK055845
12	2059.2	88.3	2174	6	ARO34241
13	1790.4	76.8	1792	9	AY007090
14	1767.4	75.8	83762	9	AP000326
15	1767.4	75.8	100000	9	AP000054
16	1767.4	75.8	100000	9	AP000169
17	1767.4	75.8	100000	17	AP000122
18	1767.4	75.8	340000	9	AP001720
19	1011.8	43.4	2216	10	CGU60263
20	803.4	34.5	2198	10	BC013551
21	798.2	34.2	2141	10	AF263239
22	755.4	32.4	798	9	AF400429
23	720	30.9	2224	10	AF263240
24	714.6	30.7	2125	10	AF260717
25	552.2	23.7	562	9	HSU53821
26	471.6	20.2	626	10	AB075973
27	470.6	20.2	597	6	AX365315
28	470.6	20.2	597	10	AF237790
29	413.8	17.8	599	6	AX365311
30	412.2	17.7	621	10	AF282255
31	403.8	17.3	597	6	AX365312
32	403.8	17.3	597	10	AF237789
33	365.2	15.7	385	6	AX156214
34	361.8	11.2	303	11	G34765
35	246.8	10.6	251	6	AX429944
36	221.6	9.5	90753	2	AC111585
37	195.2	8.4	3227	10	AB061524
38	194.4	8.3	3184	6	AX329596
39	194.4	8.3	3184	6	AX333015
40	194.4	8.3	3184	6	AX365327
41	194.4	8.3	3184	6	AX420436
42	194.4	8.3	3184	9	HUMZAKI4
43	193.2	8.3	594	6	AX365318
44	193.2	8.3	594	10	AF237791
45	192.4	8.3	659	10	AF237887

ALIGNMENTS

RESULT 1
AX365321
LOCUS AX365321 2331 bp DNA linear PAT 15-FEB-2002
DEFINITION Sequence 11 from Patent WO0204491.
ACCESSION AX365321
VERSION AX365321.1 GI:18697049
KEYWORDS human,
SOURCE Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Williams, S.R. and Rothermel, B.
TITLE Methods and compositions relating to muscle selective calcineurin
interacting protein (mcip)

JOURNAL Patent: WO 0204491-A 11 17-JAN-2002;
Board of Regents, The University of Texas System (US); Williams,
Sanders R. (US); Rothmel, Beverly (US)

FEATURES Location/Qualifiers

Source
1. 2331
/organism="Homo sapiens"
/db_xref="taxon:9606"
144. 737
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD2811.1"
/db_xref="GI:1867050"
/translation="MHFKNRYGSSSLACVANSDFSEETAKPESLPRDYKQIT
FOYFKSPKRVINSPSADARLQHKTEFLKKEKLIYPAQTLHGSHLAPPPD
KQFLISPPSPVKQVEDAPVINDLAFLSKDIPGEKTELDHAHDTTPSVVHV
CESDKEEEMERMRPKKIIQTRRPEVTPHLS"

CDS
BASE COUNT 630 a 470 c 547 g 684 t
ORIGIN

Query Match 100.0%; Score 2331; DB 6; Length 2331;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTTTTCCTCCAGGAGTGGGGGCTGGCCCTTACTGCTTATTAAGACAGCTCAAGA 60
Db 1 TTTTTCCTCCAGGAGTGGGGGCTGGCCCTTACTGCTTATTAAGACAGCTCAAGA 60
Qy 61 AGGACCTACAGCCTCTTGGAAAGAAATCTACTAGGGCTTGAAGCTGCTGTG 120
Db 61 AGGACCTACAGCCTCTTGGAAAGAAATCTACTAGGGCTTGAAGCTGCTGTG 120
Qy 121 CGCTTCTACTGTAAAGAAAGAAATGATTTTGAAGCTTAACTACAGTTTAACTCC 180
Db 121 CGCTTCTACTGTAAAGAAAGAAATGATTTTGAAGCTTAACTACAGTTTAACTCC 180
Qy 181 TGATGCTGTGTGGCAACAGTATATCTTCAGCGAAAGTGAACCAAGGCCAAATTTG 240
Db 181 TGATGCTGTGTGGCAACAGTATATCTTCAGCGAAAGTGAACCAAGGCCAAATTTG 240
Qy 241 AGTCCCTCTTTAAGACGTATGACAAAGACATCACTTCAGTATTTTAAGCTTCAAC 300
Db 241 AGTCCCTCTTTAAGACGTATGACAAAGACATCACTTCAGTATTTTAAGCTTCAAC 300
Qy 301 GAGTCAGAAATPAACTTCAGCAACCCCTTCGCCGACAGATGCGAGCTCACTGAT 360
Db 301 GAGTCAGAAATPAACTTCAGCAACCCCTTCGCCGACAGATGCGAGCTCACTGAT 360
Qy 361 AGACTGAGTTTCTGGAAAGAAATGAAGTTATTTTGTCTCAGACTTACATAGGAA 420
Db 361 AGACTGAGTTTCTGGAAAGAAATGAAGTTATTTTGTCTCAGACTTACATAGGAA 420
Qy 421 GCTCAACCTTGGCTCCGCAAAATCCAGACAGAGTTTCTGATCTCCCTCCGCTCTC 480
Db 421 GCTCAACCTTGGCTCCGCAAAATCCAGACAGAGTTTCTGATCTCCCTCCGCTCTC 480
Qy 481 CGCCAGTGGATGCAAAACAGTGAAGATGCAACCCAGTCATAACTATGATCTTTAT 540
Db 481 CGCCAGTGGATGCAAAACAGTGAAGATGCAACCCAGTCATAACTATGATCTTTAT 540
Qy 541 ATGCCATCTCCAGCTGGGGCCAGGGGAAAGTATGAATTCACAGCGAGCTACACCA 600
Db 541 ATGCCATCTCCAGCTGGGGCCAGGGGAAAGTATGAATTCACAGCGAGCTACACCA 600
Qy 601 CTCCTCAGGCTGGTGTCTCATGTATGTGAGAGTATCAAGAGAGGAAAGAGGAAA 660
Db 601 CTCCTCAGGCTGGTGTCTCATGTATGTGAGAGTATCAAGAGAGGAAAGAGGAAA 660
Qy 661 TGGAAAGATGAGAGACCTTAAGCCAAATTTATCCAGCCAGAGGCGGAGTACAGC 720
Db 661 TGGAAAGATGAGAGACCTTAAGCCAAATTTATCCAGCCAGAGGCGGAGTACAGC 720
Qy 721 CGATCACTTCACTGAACTGTGGCAGCGAGCGACTTCCAATCATACTCAAGGGA 780

Db 721 CGATCACTTCACTGAACTGTGGCAGCGAGCGAGCTTCCAATCATACTCAAGGGA 780
Qy 781 GGAATTTTACTGTGGAGGTGGCTGTGACAGCTTCTCGAGGTGGCAGCCGAGATCG 840
Db 781 GGAATTTTACTGTGGAGGTGGCTGTGACAGCTTCTCGAGGTGGCAGCCGAGATCG 840
Qy 841 GGGTGGCAGAAATCCCAAGTTGATGTTGCTCAGAGAGAAATCAAGCCGTGCTCCCTGTT 900
Db 841 GGGTGGCAGAAATCCCAAGTTGATGTTGCTCAGAGAGAAATCAAGCCGTGCTCCCTGTT 900
Qy 901 CTATGCTGCAACCAAGTTACTGTTCAATGCAACCCGGGAAATGACTTGGGCCAATCACTGA 960
Db 901 CTATGCTGCAACCAAGTTACTGTTCAATGCAACCCGGGAAATGACTTGGGCCAATCACTGA 960
Qy 961 GTTTGGTATGCAACAGCAAGCAATTTGGAGCTGTTGAGAAAACAGATAATGATAGTG 1020
Db 961 GTTTGGTATGCAACAGCAAGCAATTTGGAGCTGTTGAGAAAACAGATAATGATAGTG 1020
Qy 1021 TTTTGTACTGTTCTTTTGTGTAGGTTCTGTCTGTGCCAAGGGCAGTTGATCACTGAG 1080
Db 1021 TTTTGTACTGTTCTTTTGTGTAGGTTCTGTCTGTGCCAAGGGCAGTTGATCACTGAG 1080
Qy 1081 CTGAGAGAGAGCTTCTGTTTCTAATGAGCTTGCAGGGGCACTCTTACTGTAGGAA 1140
Db 1081 CTGAGAGAGAGCTTCTGTTTCTAATGAGCTTGCAGGGGCACTCTTACTGTAGGAA 1140
Qy 1141 GAGGTACCAAGAGAGCCGCTATGTGAGAGAGTTGTGAAAACAGACGATGCAATGT 1200
Db 1141 GAGGTACCAAGAGAGCCGCTATGTGAGAGAGTTGTGAAAACAGACGATGCAATGT 1200
Qy 1201 GGAATTTGAGAGCTTCTTCTTCCCTCAATGTTCTCATGTTGTGATATATTACTG 1260
Db 1201 GGAATTTGAGAGCTTCTTCTTCCCTCAATGTTCTCATGTTGTGATATATTACTG 1260
Qy 1261 ATTACAAAGATTAATCTTGTCTATATAAGTTACACCGTGTGTTTACATCTTTT 1320
Db 1261 ATTACAAAGATTAATCTTGTCTATATAAGTTACACCGTGTGTTTACATCTTTT 1320
Qy 1321 GGAAGCCAGAAAGCGTTTGGAAAAGTATCACTTCCCAATCTCGGATTCGTGAC 1380
Db 1321 GGAAGCCAGAAAGCGTTTGGAAAAGTATCACTTCCCAATCTCGGATTCGTGAC 1380
Qy 1381 TCTTTGCAACAGCACTTCTGCGGAACTCTCTGGAATGATCACTGAGATCCCA 1440
Db 1381 TCTTTGCAACAGCACTTCTGCGGAACTCTCTGGAATGATCACTGAGATCCCA 1440
Qy 1441 ACCGTGCAACGTGTAATGTTGTTTGGCAAAAGAGTTGATCTGAAATCTCTGTAGA 1500
Db 1441 ACCGTGCAACGTGTAATGTTGTTTGGCAAAAGAGTTGATCTGAAATCTCTGTAGA 1500
Qy 1501 ATTTAGCTTATCAATTCAGAGAAATAGCAAGTTTCACTGCCAATTTTATAGGTGAGAA 1560
Db 1501 ATTTAGCTTATCAATTCAGAGAAATAGCAAGTTTCACTGCCAATTTTATAGGTGAGAA 1560
Qy 1561 TTTTGTAGTGTGTTGGGATCGGACCTCAGTTTCTGTGTTCTTTATGTGTGTT 1620
Db 1561 TTTTGTAGTGTGTTGGGATCGGACCTCAGTTTCTGTGTTCTTTATGTGTGTT 1620
Qy 1621 TCTATATCATGATTCATAGCCAAAACTTTTGGGAAACGTGTGTGATAGTGTGTT 1680
Db 1621 TCTATATCATGATTCATAGCCAAAACTTTTGGGAAACGTGTGTGATAGTGTGTT 1680
Qy 1681 TTTTACCCAGAGACATCAAGATACACTTGTAAATTAAGCTGATAGCATATATTCATA 1740
Db 1681 TTTTACCCAGAGACATCAAGATACACTTGTAAATTAAGCTGATAGCATATATTCATA 1740
Qy 1741 CCGTGTGTAACCTTGGTGAAGAAATGATGCAAGTGGAGCTAAGATGATTAATTAACCT 1800
Db 1741 CCGTGTGTAACCTTGGTGAAGAAATGATGCAAGTGGAGCTAAGATGATTAATTAACCT 1800
Qy 1801 GTGAATCATATGTTGAGAAAGAGCTTCCCATGTCTAACAGACTTGAATTCAGAGCA 1860
Db 1801 GTGAATCATATGTTGAGAAAGAGCTTCCCATGTCTAACAGACTTGAATTCAGAGCA 1860

QY 1861 TGTCAAGTGGATAGTACTGTGGCGATATGAGGGATGAGTGCCTTTCCCATTC 1920
Db 1861 TGTCAAGTGGATAGTACTGTGGCGATATGAGGGATGAGTGCCTTTCCCATTC 1920
QY 1921 TTCTCGATGGAATGTTATATAGTAACTTGTAACTTTTCTAGTTGTAAATGTT 1980
Db 1921 TTCTCGATGGAATGTTATATAGTAACTTGTAACTTTTCTAGTTGTAAATGTT 1980
QY 1981 ATGCTGTAATAGTATATATTTGGCCCTTACAACTGCTTGAAGATGTTTATTC 2040
Db 1981 ATGCTGTAATAGTATATATTTGGCCCTTACAACTGCTTGAAGATGTTTATTC 2040
QY 2041 TGAATATCTTAATGCAAGTAACTATGCTTTTGAAGATGTTTATTC 2100
Db 2041 TGAATATCTTAATGCAAGTAACTATGCTTTTGAAGATGTTTATTC 2100
QY 2101 TTTGAGAAGCAATATGTTTGCATTAATGCTTTGATTTTCAATCAAGAAATGATTG 2160
Db 2101 TTTGAGAAGCAATATGTTTGCATTAATGCTTTGATTTTCAATCAAGAAATGATTG 2160
QY 2161 AACGTTCTCAACCCCTGTTTACGCTTGTGTAAGAGGAGCCGTTTGGGAGAGCCAT 2220
Db 2161 AACGTTCTCAACCCCTGTTTACGCTTGTGTAAGAGGAGCCGTTTGGGAGAGCCAT 2220
QY 2221 TGCATCGCTGTCGAAGTGTCTTTTGAAGTGTCTTTTAACTGAGAGGCTTCAAA 2280
Db 2221 TGCATCGCTGTCGAAGTGTCTTTTGAAGTGTCTTTTAACTGAGAGGCTTCAAA 2280
QY 2281 ATACTTTTTTAACTGCACTTATATAAATGCGCACAGATGCTCTTAC 2331
Db 2281 ATACTTTTTTAACTGCACTTATATAAATGCGCACAGATGCTCTTAC 2331

RESULT 2
HSU85267
LOCUS HSU85267 2346 bp mRNA linear PRI 19-APR-2000
DEFINITION Homo sapiens down syndrome candidate region 1 (DSCR1) gene,
alternative exon 1, complete cds.
ACCESSION U85267
VERSION U85267.2 GI:7596915
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2346)
Fuentes,J.J., Pritchard,M.A., Planas,A.M., Bosch,A., Ferrer,I. and Estivill,X.
A new human gene from the Down syndrome critical region encodes a proline-rich protein highly expressed in fetal brain and heart
Hum. Mol. Genet. 4 (10), 1935-1944 (1995)
JOURNAL MEDLINE PUBMED 96121593
REFERENCE 2 (bases 1 to 2346)
Fuentes,J.J., Pritchard,M.A. and Estivill,X.
Genomic organization, alternative splicing, and expression patterns of the DSCR1 (Down syndrome candidate region 1) gene
Genomics 44 (3), 358-361 (1997)
JOURNAL MEDLINE PUBMED 97468152
REFERENCE 3 (bases 1 to 2346)
Fuentes,J.J., Pritchard,M. and Estivill,X.
Direct Submission
Submitted (14-JAN-1997) Genetica Molecular, Institut de Recerca Oncologica, Autovia de Castelldefels, km 2.7, Hospitalet, Barcelona 08907, Spain
REFERENCE 4 (bases 1 to 2346)
Fuentes,J.J., Pritchard,M. and Estivill,X.
Direct Submission
Submitted (19-APR-2000) Genetica Molecular, Institut de Recerca Oncologica, Autovia de Castelldefels, km 2.7, Hospitalet, Barcelona 08907, Spain

REMARK Sequence update by submitter
COMMENT On Apr 19, 2000 this sequence version replaced gi:2612867.
FEATURES
Location/Qualifiers
source 1..2346
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosomes="21"
/map="21q22.1-q22.2"
gene 1..2346
/gene="DSCR1"
CDS 144..737
/gene="DSCR1"
/note="down syndrome candidate region 1; one of four alternatively spliced exon 1"
/codon_start=1
/protein_id="AAB84372.1"
/translation="MHPNPNYSFSSLIACVANSDFSESETRAKFESLFRFYDKDIT
FOYFKSFXRVIRINFSNPFSAADARLQHLKFTFLGKEMKLYFAQLHIGSSHLAPNP
KQFLISPPASPPGVKQVEDATPVINDLLYAIKLGPGKEYELHAAATDTTPSVVVHV
CESDQEEEBEEMRRPKKIIOTRPEYTPHLS"
BASE COUNT 645 a 470 c 547 g 684 t
ORIGIN
Query Match 100.0%; Score 2331; DB 9; Length 2346;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTTTTCCTTCCCGAGGAGTGGGGCTGCTTACTTCTTATAGCAGCAGCTCAAGA 60
Db 1 TTTTTCCTTCCCGAGGAGTGGGGCTGCTTACTTCTTATAGCAGCAGCTCAAGA 60
QY 61 AGGAACCTACAGCTCTTGGAAAGGAATCTCCTAGGGGCTTGACTGGGCTGTAG 120
Db 61 AGGAACCTACAGCTCTTGGAAAGGAATCTCCTAGGGGCTTGACTGGGCTGTAG 120
QY 121 CGCTTTCACTGTAAGAAAGCAGATGCAATTTTAGAACTTTTAACTACAGTTT 180
Db 121 CGCTTTCACTGTAAGAAAGCAGATGCAATTTTAGAACTTTTAACTACAGTTT 180
QY 181 TGATTGCTGTGTGGCAACAGTGTATCTTTCAGCGAAGTGAAACAGGCGCAATTT 240
Db 181 TGATTGCTGTGTGGCAACAGTGTATCTTTCAGCGAAGTGAAACAGGCGCAATTT 240
QY 241 AGTCCTCTCTTGAAGACGCTATGCAAGGACATCCTTTTCACTTATTAAGAGCTT 300
Db 241 AGTCCTCTCTTGAAGACGCTATGCAAGGACATCCTTTTCACTTATTAAGAGCTT 300
QY 301 GAGTCAGATAAATCTTCAAGCAACCCCTTCTCCGAGCAGATGCCAGGCTCCAGCT 360
Db 301 GAGTCAGATAAATCTTCAAGCAACCCCTTCTCCGAGCAGATGCCAGGCTCCAGCT 360
QY 361 AGACTGAGTTTCTGGGAAGGAATGAAGTTATATTTTGTCTCAGACCTTACACAT 420
Db 361 AGACTGAGTTTCTGGGAAGGAATGAAGTTATATTTTGTCTCAGACCTTACACAT 420
QY 421 GCTCAGACCTGCTCCGCAAAATCCAGACGAGTTTCTGATCTCCCTCCCGCTCTC 480
Db 421 GCTCAGACCTGCTCCGCAAAATCCAGACGAGTTTCTGATCTCCCTCCCGCTCTC 480
QY 481 CGCCAGTGGGATGGAACAAAGTGGAGATGCGACCCAGTCTATAAACTATGATCTT 540
Db 481 CGCCAGTGGGATGGAACAAAGTGGAGATGCGACCCAGTCTATAAACTATGATCTT 540
QY 541 ATGCCATCTCCAGCTGGGCGCAGGGAAGATGATGAATTCACGAGCAGCTGACACCA 600
Db 541 ATGCCATCTCCAGCTGGGCGCAGGGAAGATGATGAATTCACGAGCAGCTGACACCA 600
QY 601 CTCCAGCGTGGTGGTCCATGTATGTGAGAGTGTATCAAGAGAGGAGGAGAGGAAA 660
Db 601 CTCCAGCGTGGTGGTCCATGTATGTGAGAGTGTATCAAGAGAGGAGGAGAGGAAA 660
QY 661 TGGAAAGATGAGGAGACCTTAAGCCAAAAATTTATCCAGACGAGGCGCGAGTACACGC 720

Db 661 TGGAAAGAAATGAGACCTTAAGCCAAATATATCAAGCCAGAGGCGGAGATCAAGC 720
 Qy 721 CGATCCACCTCAGCTGAACTGGCAGCGGACGAGAGCGATTCGAATCATACTCAGGGA 780
 Db 721 CGATCCACCTCAGCTGAACTGGCAGCGGACGAGAGCGATTCGAATCATACTCAGGGA 780
 Qy 781 GGAATCTTTTACTGTGGAGAGTGGCTGTGTCAAGATCTTTCCGAGAGTGGCAGCGGATCG 840
 Db 781 GGAATCTTTTACTGTGGAGAGTGGCTGTGTCAAGATCTTTCCGAGAGTGGCAGCGGATCG 840
 Qy 841 GGGTGGCAGAAATCCAGTTCATGTGTCTCAGAGAGAAATCAAGGCGGTGTCCCTGTGT 900
 Db 841 GGGTGGCAGAAATCCAGTTCATGTGTCTCAGAGAGAAATCAAGGCGGTGTCCCTGTGT 900
 Qy 901 CTAAATCTGCACACAGGATTAAGTTCATGCGACCCGGGAATGACTTTGGGCCAATCATCGA 960
 Db 901 CTAAATCTGCACACAGGATTAAGTTCATGCGACCCGGGAATGACTTTGGGCCAATCATCGA 960
 Qy 961 GTTTGTGTGTGTGATGCAAGAGACATTTGGGACTGTCTGAGAAACAGATATGATAGTG 1020
 Db 961 GTTTGTGTGTGTGATGCAAGAGACATTTGGGACTGTCTGAGAAACAGATATGATAGTG 1020
 Qy 1021 TTTTGTACTGTCTTTCTGT 1080
 Db 1021 TTTTGTACTGTCTTTCTGT 1080
 Qy 1081 CTCAGAGAGAGAGCTTCTGT 1140
 Db 1081 CTCAGAGAGAGAGCTTCTGT 1140
 Qy 1141 GAGGTACCCAG 1200
 Db 1141 GAGGTACCCAG 1200
 Qy 1201 GGAATTTGTAGCGCTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 1260
 Db 1201 GGAATTTGTAGCGCTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 1260
 Qy 1261 ATTATACAGACTAATCTTTGT 1320
 Db 1261 ATTATACAGACTAATCTTTGT 1320
 Qy 1321 GGGAGAGCAG 1380
 Db 1321 GGGAGAGCAG 1380
 Qy 1381 TCTTTGCAAG 1440
 Db 1381 TCTTTGCAAG 1440
 Qy 1441 ACCGTGCAAG 1500
 Db 1441 ACCGTGCAAG 1500
 Qy 1501 ATTATAGCTTATACAAATTCAG 1560
 Db 1501 ATTATAGCTTATACAAATTCAG 1560
 Qy 1561 TTTTGTAGTTAGCTTTGTGGATTCGGAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1620
 Db 1561 TTTTGTAGTTAGCTTTGTGGATTCGGAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1620
 Qy 1621 TCTATACATGAATCATACCAAAATCTTTTGGAAATCTGTGTGTGTGTGTGTGTGTGTGTGT 1680
 Db 1621 TCTATACATGAATCATACCAAAATCTTTTGGAAATCTGTGTGTGTGTGTGTGTGTGTGTGT 1680
 Qy 1681 TTTTACCCAG 1740
 Db 1681 TTTTACCCAG 1740
 Qy 1741 CCTGTGTACACTTGTGGTGGAAAGATAGGAGTGGAGAGATGAATGAATTAACCTACTT 1800

Db 1741 CCTGTGTACACTTGTGGTGGAAAGATAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
 Qy 1801 GTGATACATATGTGTAGGAG 1860
 Db 1801 GTGATACATATGTGTAGGAG 1860
 Qy 1861 TGTCAAGTGTAGTATGATCTGTGGAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920
 Db 1861 TGTCAAGTGTAGTATGATCTGTGGAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920
 Qy 1921 TTCTGTAGTGAATTTGTATCTAGTGTAAATTTTGTATTTTGTATTTTGTATTTTGTATTTTGT 1980
 Db 1921 TTCTGTAGTGAATTTGTATCTAGTGTAAATTTTGTATTTTGTATTTTGTATTTTGTATTTTGT 1980
 Qy 1981 ATGTCTGTAAATAGTATATATTTTGGCTTACAAATCCGTAAACATGTTTGTATTTTGTATTT 2040
 Db 1981 ATGTCTGTAAATAGTATATATTTTGGCTTACAAATCCGTAAACATGTTTGTATTTTGTATTT 2040
 Qy 2041 TGAATATCTTAATGCCAAGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2100
 Db 2041 TGAATATCTTAATGCCAAGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2100
 Qy 2101 TTTGAG 2160
 Db 2101 TTTGAG 2160
 Qy 2161 AACGTTCTCAACCTGTGTAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2220
 Db 2161 AACGTTCTCAACCTGTGTAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2220
 Qy 2221 TGCATCGCTGTCCAAAGTGTCTTGTAAAGTGTCTTGTAAAGTGTCTTGTAAAGTGTCTTGTAA 2280
 Db 2221 TGCATCGCTGTCCAAAGTGTCTTGTAAAGTGTCTTGTAAAGTGTCTTGTAAAGTGTCTTGTAA 2280
 Qy 2281 ATACTTTTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2340
 Db 2281 ATACTTTTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2340
 RESULT 3
 AX281651 2348 bp DNA linear PAT 02-NOV-2001
 LOCUS AX281651
 DEFINITION Sequence 60 from Patent WO0177389.
 ACCESSION AX281651
 VERSION AX281651.1 GI:16608902
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1
 AUTHORS Shiffman, D., Somogyi, R., Lawn, R., Selkover, J. J., Porter, G. J.,
 Mikita, T. and Tai, J.
 TITLE Genes expressed in foam cell differentiation
 JOURNAL Patent: WO 0177389-A 60 18-OCT-2001;
 Incyte Genomics, Inc. (US)
 FEATURES
 source Location/Qualifiers
 1..2348
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /note="Incyte ID No: 042176-5"
 BASE COUNT 638 a 473 c 550 g 687 t
 ORIGIN
 Query Match 99.5%; Score 2318.4; DB 6; Length 2348;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 2330; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
 Qy 1 TTTTATTTTCCAG 60
 Db 8 TTTTATTTTCCAG 67
 Qy 61 AGGAACCTACAGCCTTGTGAAAGAGATCTCACTAGGAGGCTTGACTGCGTGTGTGAG 120

Db 68 AGGAACCTACAGCCTCTTGGAAGGAATCTCACTAGGGCTTGACTGGGTGGTCTGTAG 127
QY 121 CGCTTTCACTGTAGAAACAAAGATGCAATTTTAGAAACTTTAACTACAGTTTGTAGCTCCC 180
Db 128 CGCTTTCACTGTAGAAACAAAGATGCAATTTTAGAAACTTTAACTACAGTTTGTAGCTCCC 187
QY 181 TGATTTGCTGTGTGCGCAACAGTATATCTTCAGCGAAAGTGAACACAGGCGCCAAATTTG 240
Db 188 TGATTTGCTGTGTGCGCAACAGTATATCTTCAGCGAAAGTGAACACAGGCGCCAAATTTG 247
QY 241 AGTCCCTCTTTAGGACGTATGCAAGGACATCACCTTTTCAGTATTTTAAAGACTTCAAAAC 300
Db 248 AGTCCCTCTTTAGGACGTATGCAAGGACATCACCTTTTCAGTATTTTAAAGACTTCAAAAC 307
QY 301 GAGTCAGAAATAACTTTCAGCAACCCCTTCTCCGACAGAGATGCCAGGCTCCAGCTGCATA 360
Db 308 GAGTCAGAAATAACTTTCAGCAACCCCTTCTCCGACAGAGATGCCAGGCTCCAGCTGCATA 367
QY 361 AGACTGAGTTTCTCGGAAGGAATGAAGTTATATTTTGTCTCAGACCTTACATAGGAA 420
Db 368 AGACTGAGTTTCTCGGAAGGAATGAAGTTATATTTTGTCTCAGACCTTACATAGGAA 427
QY 421 GCTCACACT-GGCTCCGCCAAATCCAGACAAGCAGTTTCTGATCTCCCTCCCGCCTCT 479
Db 428 GCTCACACTGGGCTCCGCCAAATCCAGACAAGCAGTTTCTGATCTCCCTCCCGCCTCT 487
QY 480 CGCCAGTGGGATGGAAACAAGTGGAGATGCGACCCAGTCATAAACTATGATCTCTTA 539
Db 488 CGCCAGTGGGATGGAAACAAGTGGAGATGCGACCCAGTCATAAACTATGATCTCTTA 547
QY 540 TATGCCATCTCCAAGCTGGGCGAGGGAAGATGAATTTGATGCGAGCGAGTCAACACC 599
Db 548 TATGCCATCTCCAAGCTGGGCGAGGGAAGATGAATTTGATGCGAGCGAGTCAACACC 607
QY 600 ACTCCAGCTGGTGGTCCATGTATGTGAGAGTGATCAAGAGAGGAGGAAGAGGAA 659
Db 608 ACTCCAGCTGGTGGTCCATGTATGTGAGAGTGATCAAGAGAGGAGGAAGAGGAA 667
QY 660 ATGGAAGAATGAGGAGACCTTAAGCCAAAATTTATCCAGACCAGGAGCGGAGTACACG 719
Db 668 ATGGAAGAATGAGGAGACCTTAAGCCAAAATTTATCCAGACCAGGAGCGGAGTACACG 727
QY 720 CCGATCCACCTCAGCTGAACCTGGCACCGCAGGAGCGCATTCGAATCATCTACAGGG 779
Db 728 CCGATCCACCTCAGCTGAACCTGGCACCGCAGGAGCGCATTCGAATCATCTACAGGG 787
QY 780 AGGAATCTTTTACTGTGGAGGTGGCTGGTCAACGACTTCTTCGGAGGTGGCAGCGGAGATC 839
Db 788 AGGAATCTTTTACTGTGGAGGTGGCTGGTCAACGACTTCTTCGGAGGTGGCAGCGGAGATC 847
QY 840 GGGTGGCAGAAATCCCAAGTTTCATGTGCTCAGAGAGAAATCAAGCGCGTGTCCCTTGT 899
Db 848 GGGTGGCAGAAATCCCAAGTTTCATGTGCTCAGAGAGAAATCAAGCGCGTGTCCCTTGT 907
QY 900 TCTAATGCTGCACACCACTGTTCTATGGCACCCGGGAATGACATTTGGCCCAATCACTG 959
Db 908 TCTAATGCTGCACACCACTGTTCTATGGCACCCGGGAATGACATTTGGCCCAATCACTG 967
QY 960 AGTTTGTGGTATGCAACAAGGACATTTGGGACTGTCTTGAGAAAACAGATAATGATGT 1019
Db 968 AGTTTGTGGTATGCAACAAGGACATTTGGGACTGTCTTGAGAAAACAGATAATGATGT 1027
QY 1020 GTTTTGTACTGTTCTTTCTGCTAGGTTCTGCTGCTGCAAGGGCAGGTTGATCAGTGA 1079
Db 1028 GTTTTGTACTGTTCTTTCTGCTAGGTTCTGCTGCTGCAAGGGCAGGTTGATCAGTGA 1087
QY 1080 GCTCAGAGAGAGCTTCTGTTTCTAAGTGGCTTCAGGGGCACTCTCTACTGTTAGGA 1139
Db 1088 GCTCAGAGAGAGCTTCTGTTTCTAAGTGGCTTCAGGGGCACTCTCTACTGTTAGGA 1147
QY 1140 AGAGGTACCAAGGAGCCCTAGTGGAGAGAGGTTGTGAAAACAGCAGCAATGCAATG 1199

Db 1148 AGAGTACCACAGGAAGCCCTAGTGCAGAGAGGTTGTGAAAAACAGCAGCAATGCAATG 1207
QY 1200 TGGAAATTTGAGCGTTTCTTTCTTCCCTCATGTTCTCATGTTTGTGCATGTATATTAAT 1259
Db 1208 TGGAAATTTGAGCGTTTCTTTCTTCCCTCATGTTCTCATGTTTGTGCATGTATATTAAT 1267
QY 1260 GATTTTACAAGACTAAACCTTTGTTGCTGATATAAAGTTTACACCGTTGTTGTGTACATCTTT 1319
Db 1268 GATTTTACAAGACTAAACCTTTGTTGCTGATATAAAGTTTACACCGTTGTTGTGTACATCTTT 1327
QY 1320 TGGGAAGCCAGGAAGCGTTTGGAAAAACGTTATCACCTTTTCCCAAGATTTCTCGGATTTCTCGA 1379
Db 1328 TGGGAAGCCAGGAAGCGTTTGGAAAAACGTTATCACCTTTTCCCAAGATTTCTCGGATTTCTCGA 1387
QY 1380 CTCTTTGCCAACAGCAGCTTGTGCGGAACTCTTCTCGGAATGCATTTCACTCAGCAATCCCC 1439
Db 1388 CTCTTTGCCAACAGCAGCTTGTGCGGAACTCTTCTCGGAATGCATTTCACTCAGCAATCCCC 1447
QY 1440 AACCGTGCACACGTGTAACTTTGCTTTTGCAAAAAGAGTTGATCTGAAATTTCTCTGTAG 1499
Db 1448 AACCGTGCACACGTGTAACTTTGCTTTTGCAAAAAGAGTTGATCTGAAATTTCTCTGTAG 1507
QY 1500 AATTAGCTTTATACAAATTCAGAGAATAGCAGTTTCACTGCCAACTTTTAGTCGGTGAGAA 1559
Db 1508 AATTAGCTTTATACAAATTCAGAGAATAGCAGTTTCACTGCCAACTTTTAGTCGGTGAGAA 1567
QY 1560 ATTTTAGTTTGTGTTGGGATCGGACCTCAGTTTCTGTTGTTTCTTTTATGTGGTGT 1619
Db 1568 ATTTTAGTTTGTGTTGGGATCGGACCTCAGTTTCTGTTGTTTCTTTTATGTGGTGT 1627
QY 1620 TTCTATACATGAATCATAGCCAAAACTTTTGGAAAACTTTTGGAACTGTGTGGTGTGAGTAGTGGT 1679
Db 1628 TTCTATACATGAATCATAGCCAAAACTTTTGGAAAACTTTTGGAACTGTGTGGTGTGAGTAGTGGT 1687
QY 1680 CTTTTACCCCGAAGACATCAAGATACTTGTAAATAAAGCTGATAGCATATATTCAT 1739
Db 1688 CTTTTACCCCGAAGACATCAAGATACTTGTAAATAAAGCTGATAGCATATATTCAT 1747
QY 1740 ACTGTTGTACACTTTGGGTGAAAGTATGCGAGTGGGAGACTAAGATGTATTAACCTACC 1799
Db 1748 ACTGTTGTACACTTTGGGTGAAAGTATGCGAGTGGGAGACTAAGATGTATTAACCTACC 1807
QY 1800 TGTGAATCATATGTTGTAGGAAAAAGCTGTTCCCATGTCTAAACAGACTTGAATTTCAAAGC 1859
Db 1808 TGTGAATCATATGTTGTAGGAAAAAGCTGTTCCCATGTCTAAACAGACTTGAATTTCAAAGC 1867
QY 1860 ATGTCAGTGTAGTAGTAGTCTGTGGCGATATGAGAGGATGCGAGTGCCTTTTCCCAATTC 1919
Db 1868 ATGTCAGTGTAGTAGTAGTCTGTGGCGATATGAGAGGATGCGAGTGCCTTTTCCCAATTC 1927
QY 1920 ATTCTGATGGAATTTGTATATACTAGGTTTAACTTTGTAATTTTCTTAGTTGTAATGTG 1979
Db 1928 ATTCTGATGGAATTTGTATATACTAGGTTTAACTTTGTAATTTTCTTAGTTGTAATGTG 1987
QY 1980 TATGTCGTGTAATAGGTATTTATATTTTGGCTTTCAATACCGTAAACAATGTTGTCAAT 2039
Db 1988 TATGTCGTGTAATAGGTATTTATATTTTGGCTTTCAATACCGTAAACAATGTTGTCAAT 2047
QY 2040 TTGAAATACCTTAATGCCAAGTAACTGCAATGCTTTGGAATTTTGGAGATGTTTAT 2099
Db 2048 TTGAAATACCTTAATGCCAAGTAACTGCAATGCTTTGGAATTTTGGAGATGTTTAT 2107
QY 2100 CTTTGTAGAAGCAAAATATGTTTGTGCAATTAATGCTTTGATATGTCATATCAAGAAATTCAT 2159
Db 2108 CTTTGTAGAAGCAAAATATGTTTGTGCAATTAATGCTTTGATTTTATATCAAGAAATTCAT 2167
QY 2160 GAACGTTTCTCAAACCTGTTTACGGTACTTGGTAAGAGGAGCGGTTTGGGAGAGACCA 2219
Db 2168 GAACGTTTCTCAAACCTGTTTACGGTACTTGGTAAGAGGAGCGGTTTGGGAGAGACCA 2227
QY 2220 TTGCATCGCTGCCAAGTGTCTTGTAGTGTCTTTTAACTGGAGAGGCTAACCTCAA 2279
Db 2228 TTGCATCGCTGCCAAGTGTCTTGTAGTGTCTTTTAACTGGAGAGGCTAACCTCAA 2287

QY 2280 AATACTTTTAACTGATCTCTATTAATTAATGGGACAGATAGTCTCTTAC 2331
 Db 2288 AATATTTTAACTGATCTCTATTAATTAATGGGACAGATAGTCTCTTAC 2339
 RESULT 4
 AX365324 2212 bp DNA linear PAT 15-FEB-2002
 LOCUS Sequence 14 from Patent WO204491.
 DEFINITION AX365324
 ACCESSION AX365324.1 GI:18697051
 VERSION
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1
 AUTHORS Williams, S.R. and Rochemel, B.
 TITLE Methods and compositions relating to muscle selective calcineurin
 JOURNAL interacting protein (mciip)
 Board of Regents, The University of Texas System (US) ; Williams,
 Sanders R. (US) ; Rochemel, Beverly (US)
 FEATURES
 source location/Qualifiers
 1..2212
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 25..618
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD23812.1"
 /db_xref="GI:18697052"
 /translation="MBEVDLQDLSATTAACHLDPVDFVDCRAKFSLEPTDNDIT
 FQYFKSFKRVIRFNSPSPADARLQHLKTEFLGEMULYPAQTHISSHLAPND
 KQFLSPSPSPVPMKQVEDATPVINYLDAISLKGEGEKELHAADITPSVYVHV
 CESDQKEEEMEMRRPKPIQTRREYPIHLS"
 BASE COUNT 588 a 455 c 529 g 640 t
 ORIGIN
 Query Match 90.5%; Score 2109; DB 6; Length 2212;
 Best Local Similarity 97.5%; Pred. No. 0;
 Matches 2142; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
 QY 135 GAAAGCAAGTGAATTTTAACTTAACAGTTTGAAGTCCCTGATGCTGTG 194
 Db 16 GACTGCGAGATGGAGGAGTGAGCTGACAGACCTGCCAGCCGACCACTGCTGAC 75
 QY 195 GCAACAGTATATCTTCAAGCAAGTGAACCCAGGCCAAATTTGAGTCCCTTTAGG 254
 Db 76 CTGAGCCCGGGCGTGTGCGAGAGCGCTGTGCGGGCCAAATTTGAGTCCCTTTAGG 135
 QY 255 ACGTATGACAGACATCACTTTCATATTTTAAAGCTTCAAGAGTCAAGATTAAC 314
 Db 136 ACGTATGACAGACATCACTTTCATATTTTAAAGCTTCAAGAGTCAAGATTAAC 195
 QY 315 TTGAGCAACCCCTCTCGAGCAGATGCCAGGCTCCAGCTGATTAAGTGAAGTTTCTG 374
 Db 196 TTGAGCAACCCCTCTCGAGCAGATGCCAGGCTCCAGCTGATTAAGTGAAGTTTCTG 255
 QY 375 GAAAGCAAGTGAATTTTGTCTGAGCTTCAACATAGAGAGTCAACCTGGCT 434
 Db 256 GAAAGCAAGTGAATTTTGTCTGAGCTTCAACATAGAGAGTCAACCTGGCT 315
 QY 435 CCGGCAATTCAGACAGATTTCTGATCTCCCTCCGCTCTCGGCGAGTGGATGG 494
 Db 316 CCGGCAATTCAGACAGATTTCTGATCTCCCTCCGCTCTCGGCGAGTGGATGG 375
 QY 495 AAACAAGTGAAGATGCGACCCAGCTCATTAAGTATCTTATATGCGATCTCAAG 554
 Db 376 AAACAAGTGAAGATGCGACCCAGCTCATTAAGTATCTTATATGCGATCTCAAG 435
 QY 555 CTGGGCGCAGGGGAAAGTATGAATTGCAAGCAGCAGTCAACACCTCCAGGCTGTG 614

Db 436 CTGGGCGCAGGGGAAAGTATGAATTGCAAGCAGCAGTCAACCACTCCAGCGTGTG 495
 QY 615 GTCCATGATATGAGAGTGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 674
 Db 496 GTCCATGATATGAGAGTGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 555
 QY 675 AGACTTAAGCCAAATTAATCCAGACAGAGAGCCGAGTACACGCCATCCACTCAGC 734
 Db 556 AGACTTAAGCCAAATTAATCCAGACAGAGAGCCGAGTACACGCCATCCACTCAGC 615
 QY 735 TGAATGCGACCGGACAGAGAGCCCATTCCTCAATCTCAAGGAGAGATCTTTACTG 794
 Db 616 TGAATGCGACCGGACAGAGAGCCCATTCCTCAATCTCAAGGAGAGATCTTTACTG 675
 QY 795 TGAAGTGGCTGTGTCAGCATCTTCTCGAGAGTGGCAGCCGATTCGGGTGGCAGAAATC 854
 Db 676 TGAAGTGGCTGTGTCAGCATCTTCTCGAGAGTGGCAGCCGATTCGGGTGGCAGAAATC 735
 QY 855 CCAGTTCATGTTGCTCAGAGAGAGATCAAGCCGCTGCTCCCTGTTCTAATGCTGCACAC 914
 Db 736 CCAGTTCATGTTGCTCAGAGAGAGATCAAGCCGCTGCTCCCTGTTCTAATGCTGCACAC 795
 QY 915 CAGTTACTGTTCAAGCAGACCCGGGAATGACTGGGCAATCACTGAGTTGTGTGATCG 974
 Db 796 CAGTTACTGTTCAAGCAGACCCGGGAATGACTGGGCAATCACTGAGTTGTGTGATCG 855
 QY 975 CACAAGACATTTGGAGCTGTCTTGAGAAAAAGATATATATATATATATATATATAT 1034
 Db 856 CACAAGACATTTGGAGCTGTCTTGAGAAAAAGATATATATATATATATATATATAT 915
 QY 1035 TTTTCTGTTAGTCTCTGTCTGTCAGAGGCGAGTTATATAGAGTTCAGGAGAGAGCT 1094
 Db 916 TTTTCTGTTAGTCTCTGTCTGTCAGAGGCGAGTTATATATATATATATATATATAT 975
 QY 1095 TCCGTTTCTAAGTGGCCCTGACAGGCGCACTCTACTGTAGTGAAGAGTATCACAGAGA 1154
 Db 976 TCCGTTTCTAAGTGGCCCTGACAGGCGCACTCTACTGTAGTGAAGAGTATCACAGAGA 1035
 QY 1155 AGCCGCTAGTGCAGAGAGTGTGAAAAACAGACAGCAATGCATGTGAAATTTAGCCT 1214
 Db 1036 AGCCGCTAGTGCAGAGAGTGTGAAAAACAGACAGCAATGCATGTGAAATTTAGCCT 1095
 QY 1215 TTCCCTTTCTCCCTCAGATTCCTGATTTTGGAGTATATATATATATATATATATAT 1274
 Db 1096 TTCCCTTTCTCCCTCAGATTCCTGATTTTGGAGTATATATATATATATATATATAT 1155
 QY 1275 CCTTGTGTTAT 1334
 Db 1156 CCTTGTGTTAT 1215
 QY 1335 GCGTTTGGAAAAAGTATACCTTTTCCAGATTTCTGGAATTCCTGACCTTTTGGACAGACA 1394
 Db 1216 GCGTTTGGAAAAAGTATACCTTTTCCAGATTTCTGGAATTCCTGACCTTTTGGACAGACA 1275
 QY 1395 CTTCGTTGGGAACTCTTCTCGGAAATGCACTCAGATCCCAACCGTGAACGCTGT 1454
 Db 1276 CTTCGTTGGGAACTCTTCTCGGAAATGCACTCAGATCCCAACCGTGAACGCTGT 1335
 QY 1455 AACTGTGCTTTTGGAAAAAGATGATCTGAATTTCTCTGTAGAAATTTAGCTTATACA 1514
 Db 1336 AACTGTGCTTTTGGAAAAAGATGATCTGAATTTCTCTGTAGAAATTTAGCTTATACA 1395
 QY 1515 ATTCAAGAAATGCAATTTCACTGCCAATTTTATGGGTGAGAAATTTTATTTAGGTG 1574
 Db 1396 ATTCAAGAAATGCAATTTCACTGCCAATTTTATGGGTGAGAAATTTTATTTAGGTG 1455
 QY 1575 TTTGGATGAGACCTCAGTTTCTGTTGTTCTTTATATGAGTGTCTTATATATATATAT 1634
 Db 1456 TTTGGATGAGACCTCAGTTTCTGTTGTTCTTTATATGAGTGTCTTATATATATATAT 1515
 QY 1635 ATAGCCAAATTTTGGAAATCTGTTGTTGAGATGTTGTTTACCCACGAA 1694

Db 1516 ATAGCCAAAAAAGCTTTTGGAAAGCTTGGTTGAGATAGTTGGTTCTTTTACCCACAGAA 1575
QY 1695 GACATCAAGATACACTTGTAAATAAGCTGATAGCATATATTCATACACTTGTGTACACTT 1754
Db 1576 GACATCAAGATACACTTGTAAATAAGCTGATAGCATATATTCATACACTTGTGTACACTT 1635
QY 1755 GGGTGAAGATAGTGGCAGTGGGAGACTAGATGATTAACCTACCTGTGAATCATATGTT 1814
Db 1636 GGGTGAAGATAGTGGCAGTGGGAGACTAGATGATTAACCTACCTGTGAATCATATGTT 1695
QY 1815 GTAGGAAGAGCTGTTCCCATGTTCTAACAGGACTTGAATTTCAAAGCATGCTCAAGTGGATAG 1874
Db 1696 GTAGGAAGAGCTGTTCCCATGTTCTAACAGGACTTGAATTTCAAAGCATGCTCAAGTGGATAG 1755
QY 1875 TAGATCTGTGGGATATGAGAGGATGAGTGCCTTTCCCATTCATTCTCTGATGGAATT 1934
Db 1756 TAGATCTGTGGGATATGAGAGGATGAGTGCCTTTCCCATTCATTCTCTGATGGAATT 1815
QY 1935 GTTATCTAGGTTAACTTTGTAATTTTCTAGTTGTAATGTTGTAATGTTGTAATGTTGTAATG 1994
Db 1816 GTTATCTAGGTTAACTTTGTAATTTTCTAGTTGTAATGTTGTAATGTTGTAATGTTGTAATG 1875
QY 1995 GGTATTATATTTGGCCTTACAACTACCTGAACAAATGTTTGTCAATTTGAAATCTTAATG 2054
Db 1876 GGTATTATATTTGGCCTTACAACTACCTGAACAAATGTTTGTCAATTTGAAATCTTAATG 1935
QY 2055 CCAAGTAACTAATGATGCTTTGGAATTTGGAAGATGTTTATCTTTGAGAAGCAAT 2114
Db 1936 CCAAGTAACTAATGATGCTTTGGAATTTGGAAGATGTTTATCTTTGAGAAGCAAT 1995
QY 2115 ATGTTTCAATTAATGCTTTGATTTGTTCAATCAAGAAATGATTGAACCTTCTCAAC 2174
Db 1996 ATGTTTCAATTAATGCTTTGATTTGTTCAATCAAGAAATGATTGAACCTTCTCAAC 2055
QY 2175 CTGTTTACGTTACTGTTGAAGAGGAGCCGTTTGGAGAGACCAATGTCATGCTGTCCA 2234
Db 2056 CTGTTTACGTTACTGTTGAAGAGGAGCCGTTTGGAGAGACCAATGTCATGCTGTCCA 2115
QY 2235 AGTGTCTTCTGTTAAAGTCTTTTAACTGGAGAGCTTAACTCAAAATCTTTTAACT 2294
Db 2116 AGTGTCTTCTGTTAAAGTCTTTTAACTGGAGAGCTTAACTCAAAATCTTTTAACT 2175
QY 2295 TGCATCTATAATAATGGGCACAGTATGCTCTTAC 2331
Db 2176 TGCATCTATAATAATGGGCACAGTATGCTCTTAC 2212

RESULT 5
LOCUS HSU85266 2227 bp mRNA linear PRI 19-APR-2000
DEFINITION Homo sapiens down syndrome candidate region 1 (DSCR1) gene,
alternative exon 1, complete cds.
ACCESSION U85266
VERSION U85266.2 GI:7596913
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2227)
AUTHORS Fuentet, J.J., Pritchard, M.A., Planas, A.M., Bosch, A., Ferrer, I. and
Estivill, X.
TITLE A new human gene from the Down syndrome critical region encodes a
proline-rich protein highly expressed in fetal brain and heart
JOURNAL Hum. Mol. Genet. 4 (10), 1935-1944 (1995)
MEDLINE 96121593
PUBMED 8595418
REFERENCE 2 (bases 1 to 2227)
AUTHORS Fuentet, J.J., Pritchard, M.A. and Estivill, X.
TITLE Genomic organization, alternative splicing, and expression patterns
of the DSCR1 (Down syndrome candidate region 1) gene
JOURNAL Genomics 44 (3), 358-361 (1997)
MEDLINE 97468152

PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

9325060
3 (bases 1 to 2227)
Fuentes, J.J., Pritchard, M. and Estivill, X.
Direct Submission
Submitted (14-JAN-1997) Genetica Molecular, Institut de Recerca
Oncologica, Autovia de Castelldefels, km 2.7, Hospitalet, Barcelona
08907, Spain

REFERENCE
AUTHORS
TITLE
JOURNAL

4 (bases 1 to 2227)
Fuentes, J.J., Pritchard, M. and Estivill, X.
Direct Submission
Submitted (19-APR-2000) Genetica Molecular, Institut de Recerca
Oncologica, Autovia de Castelldefels, km 2.7, Hospitalet, Barcelona
08907, Spain

REMARK

Sequence update by submitter

COMMENT

On Apr 19, 2000 this sequence version replaced gi:2612865.
Fuentes, J.J. Hum.Mol.Genet. 4, 1935-1944, 1995.

FEATURES

source

1..2227
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.1-q22.2"
1..2227
/gene="DSCR1"
25..618
/gene="DSCR1"
/note="down syndrome candidate region 1; one of four
alternatively spliced exon 1"
/codon_start=1
/protein_id="AA84371.2"
/db_xref="GI:7596914"
/translation="MEVDLQDLPSATIACHLDPRVFDGLCRKFSLEPTVDKQIT
FOYFKPKRVIRNPSNPSAARDLQHLKTEFLGKEMKLYFAQTLHGSHLAPPNDP
KQFLLSPAPPPVGVKWDTPVINTDLYIAISKLGPBGKYLHAATDTTPSVVHV
CESQDEKEBEEMRRPKPIQTRRPEYTHLS"

gene

CDS

603 a 455 c 529 g 640 t

Query Match

Best Local Similarity 90.5%; Score 2109; DB 9; Length 2227;

Matches 2142; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 135 GAAAGCAAGATGCACTTTTAGAACTTTAACTACAGTTTAACTCCCTGATTCCTCTGTGTG 194

Db 16 GACTGCGAGATGGAGGAGTGGACCTGCAGGACCTGCCAGGCCACCATCGCTGTAC 75

QY 195 GCAACAGTGTATCTTCAGCGAAGTGAACACGAGGCAAAATTTAGTCCCTCTTAGG 254

Db 76 CTGACCCGCGCGTGTTCGTGGACGCGCTGTGCGGGCCAAATTTAGTCCCTCTTAGG 135

QY 255 ACATATGACAGGACATCACCTTTTCACTATTTTAAAGAGCTTCAACAGGATCAGATTAAC 314

Db 136 ACATATGACAGGACATCACCTTTTCACTATTTTAAAGAGCTTCAACAGGATCAGATTAAC 195

QY 315 TTACGCAACCCCTTCTCGGAGCAGATGCCAGGCTCCAGCTCCAGTGAAGAGTCTTCTG 374

Db 196 TTACGCAACCCCTTCTCGGAGCAGATGCCAGGCTCCAGCTCCAGTGAAGAGTCTTCTG 255

QY 375 GGAAGGAAATGAGTTATATTTTGTCTCAGACCTTACATAGAGAGTCCACACCTGGCT 434

Db 256 GGAAGGAAATGAGTTATATTTTGTCTCAGACCTTACATAGAGAGTCCACACCTGGCT 315

QY 435 CCGCAAAATCCAGACAGAGTCTTCTGATCTCCCTCCGCGCTCTCGGCGAGTGGATGG 494

Db 316 CCGCAAAATCCAGACAGAGTCTTCTGATCTCCCTCCGCGCTCTCGGCGAGTGGATGG 375

QY 495 AACCAAGTGAAGATGGACCCGAGTCAATAAATATGATCTTATATGCCATCTCCAAG 554

Db 376 AACCAAGTGAAGATGGACCCGAGTCAATAAATATGATCTTATATGCCATCTCCAAG 435

QY 555 CTGGGCGCAGGGGAAAGTGAATTTGCGGAGGAGTGAACCACTCCACGCTGGTG 614

Db 436 CTGGGCGCAGGGGAAAGTGAATTTGCGGAGGAGTGAACCACTCCACGCTGGTG 495

615 GTCCATGATGTGAGAGTATCAAGAGAAAGAGAAAGAAATGGAAGAATGAG 674
 496 GTCCATGATGTGAGAGTATCAAGAGAAAGAGAAAGAAATGGAAGAATGAG 555
 675 AGACCTTAAGCCAAAATATTCAGACAGAGAGCCGAGTACACGCCATCCACTCAGC 734
 556 AGACCTTAAGCCAAAATATTCAGACAGAGAGCCGAGTACACGCCATCCACTCAGC 615
 725 TGAACCTGAGCAGCAGAGAGAGATTCCTCAATCACTACAGGAGAAATCTTTTACTG 794
 616 TGAACCTGAGCAGCAGAGAGAGATTCCTCAATCACTACAGGAGAAATCTTTTACTG 675
 795 TGAAGTGTGCTGTCAAGCTTCTTCGAGAGTGGAGAGCCGAGATGGGGTGGCAGAAATC 854
 676 TGAAGTGTGCTGTCAAGCTTCTTCGAGAGTGGAGAGCCGAGATGGGGTGGCAGAAATC 735
 855 CCAGTTCATGTTGCTCAGAGAGAAATCAAGGCCGTGTCCCTTGTCTTAATGCTGCAGAC 914
 736 CCAGTTCATGTTGCTCAGAGAGAAATCAAGGCCGTGTCCCTTGTCTTAATGCTGCAGAC 795
 915 CAGTTACTGTTCATGAGCAGCCGAGAAATGACTTGGCCAACTACTGAGTTGTGTGATCG 974
 796 CAGTTACTGTTCATGAGCAGCCGAGAAATGACTTGGCCAACTACTGAGTTGTGTGATCG 855
 975 CACAGAGACATTTGGGACTGTCTTGAGAAAACAGATAGATAGTTTGTACTGTTC 1034
 856 CACAGAGACATTTGGGACTGTCTTGAGAAAACAGATAGATAGTTTGTACTGTTC 915
 1035 TTTTCTGTGAGTCTGTCTGTCTGCAGAGGCGAGTTGATCACTAGCTCAGAGAGAGCT 1094
 916 TTTTCTGTGAGTCTGTCTGTCTGCAGAGGCGAGTTGATCACTAGCTCAGAGAGAGCT 975
 1095 TCTGTCTTAAGTGGCTCTCAGAGGCGCACTCTCTACTGTTAGAGAGAGTACCAAGCA 1154
 976 TCTGTCTTAAGTGGCTCTCAGAGGCGCACTCTCTACTGTTAGAGAGAGTACCAAGCA 1035
 1155 AGCGGCTTAAGTGGAGAGGTTGAAAACAGCAGCAATCAATGTGAAATTTGACCGT 1214
 1036 AGCGGCTTAAGTGGAGAGGTTGAAAACAGCAGCAATCAATGTGAAATTTGACCGT 1095
 1215 TTCTTTCTTCCCTCATGTCTCATGTTTGTGATGATATTAATCTGATTTTACAAAGCTAA 1274
 1096 TTCTTTCTTCCCTCATGTCTCATGTTTGTGATGATATTAATCTGATTTTACAAAGCTAA 1155
 1275 CTTTGTTCTGATATTAAGTATACACCGTGTGTTTAACTTTTGGAGACAGAGAA 1334
 1156 CTTTGTCTGATATTAAGTATACACCGTGTGTTTAACTTTTGGAGACAGAGAA 1215
 1335 GCGTTTGGAAAAGTATCACCTTCCAGATTCGAGATTCCTGACTCTTTGGACAGCA 1394
 1216 GCGTTTGGAAAAGTATCACCTTCCAGATTCGAGATTCCTGACTCTTTGGACAGCA 1275
 1395 CTTCCTTCGCGAACTCTTCTCGAATGCAATCACTCAGCATCCCAACCGTGCAACGTGT 1454
 1276 CTTCCTTCGCGAACTCTTCTCGAATGCAATCACTCAGCATCCCAACCGTGCAACGTGT 1335
 1455 AACTTGTCTTTTGGAAAAGTATGAAATTCCTCTGTAAGATTTAGCTTATACA 1514
 1336 AACTTGTCTTTTGGAAAAGTATGAAATTCCTCTGTAAGATTTAGCTTATACA 1395
 1515 ATTGAGAGATAGCAGTTTCACTGCAACTTTAGTGGGTGAGAAATTTAGTTAGGTG 1574
 1336 ATTGAGAGATAGCAGTTTCACTGCAACTTTAGTGGGTGAGAAATTTAGTTAGGTG 1455
 1575 TTTGGGATCGAGCTCACTTCTGTGTTTCTTTTATGTGTGTTCTAATACATGATC 1634
 1486 TTTGGGATCGAGCTCACTTCTGTGTTTCTTTTATGTGTGTTCTAATACATGATC 1515
 1635 ATAGCAAAAACCTTTTGGAAAAGTGTGTTGAGATGTTGTTCTTTTACCCACAGAA 1694
 1516 ATAGCAAAAACCTTTTGGAAAAGTGTGTTGAGATGTTGTTCTTTTACCCACAGAA 1575

1695 GACATCAAGATACCTTGTAAATTAAGCTGATAGCATATATTCATCTGTGTACCTT 1754
 1576 GACATCAAGATACCTTGTAAATTAAGCTGATAGCATATATTCATCTGTGTACCTT 1635
 1755 GGGTGAAGAGTATGAGAGAGGAGACTAAGATGATTAATCACTACCTGTAATGTT 1814
 1636 GGGTGAAGAGTATGAGAGAGGAGACTAAGATGATTAATCACTACCTGTAATGTT 1695
 1815 GTAGAAAAGCTGTCCCATATGCTTAAACAGACTGAATCAAGCATGCAAGTGATG 1874
 1696 GTAGAAAAGCTGTCCCATATGCTTAAACAGACTGAATCAAGCATGCAAGTGATG 1755
 1875 TAGATCTGTGCGATATGAGAGGATGCACTGCTTCCCATATGCTGATGGAAT 1934
 1756 TAGATCTGTGCGATATGAGAGGATGCACTGCTTCCCATATGCTGATGGAAT 1815
 1935 GTTATCTAGGTTAACTTGTGATTTTCTTCTAGTTGTAATGTATGCTGTGAAATA 1994
 1816 GTTATCTAGGTTAACTTGTGATTTTCTTCTAGTTGTAATGTATGCTGTGAAATA 1875
 1995 GGTATTAATTTTGGCCCTTAAACATACCGTAAACATGTTGTCAATTTGAAATCTTAATG 2054
 1876 GGTATTAATTTTGGCCCTTAAACATACCGTAAACATGTTGTCAATTTGAAATCTTAATG 1935
 2055 CCAAGTAAACAATGATGCTTTGGAAAATTTGGAAATGCTTTTATCTTGAGAGCAAT 2114
 1936 CCAAGTAAACAATGATGCTTTGGAAAATTTGGAAATGCTTTTATCTTGAGAGCAAT 1995
 2115 ATGTTTGATTAATGCTTTGATGTTGATATCATATCAAGAAATGTAATGTAACGTTCAACC 2174
 1996 ATGTTTGATTAATGCTTTGATGTTGATATCATATCAAGAAATGTAATGTAACGTTCAACC 2055
 2175 CTGTTACGTAATGTTGTAAGAGGAGCCGTTTGGAGAGACATTCGCTGTCCA 2234
 2056 CTGTTACGTAATGTTGTAAGAGGAGCCGTTTGGAGAGACATTCGCTGTCCA 2115
 2235 AGTGTCTTCTTAAAGCTTTTAACTGAGAGAGCTAACCTCAAAATYACTTTTAAAC 2294
 2116 AGTGTCTTCTTAAAGCTTTTAACTGAGAGAGCTAACCTCAAAATYACTTTTAAAC 2175
 2295 TGCATTTCTAATAAATGAGCAGATATGCTCTTAC 2331
 2176 TGCATTTCTAATAAATGAGCAGATATGCTCTTAC 2212

RESULT 6
 BC002864 2289 bp mRNA linear PRI 12-JUL-2001
 LOCUS
 DEFINITION Homo sapiens, Down syndrome critical region gene 1, clone MGC:10263
 IMAGE:3944959, mRNA, complete cds.
 ACCESSION
 BC002864
 VERSION
 BC002864.1 GI:12804022
 KEYWORDS
 MGC.
 SOURCE
 Homo sapiens.
 ORGANISM
 Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 2289)
 REFERENCE
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (05-FEB-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 REMARK
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 COMMENT
 Contact: MGC help desk
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: DCDT/DRP
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland;

Web site: <http://www.nisc.nih.gov/>
 Contact: nisc_mgcn@hgr1.nih.gov

Shvchenko, Y., Wecherby, K.D., Beckstrom-Sternberg, S.M.,
 Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,
 Lim, M., Maduro, O.L., Masiello, C., Mastrian, S.D., McLooskey, J.C.,
 McDowell, J., Pearson, S., Snyder, B., Stantripop, S., Thomas, P.J.,
 Tionsgon, E., Touchman, J.W., Tsurgueon, C., Vogt, J.L., Walker, M.A.,
 Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IPAL Plate: 15 Row: 4 Column: 23
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA qi: 7596913.

source

```

location/Qualifiers
1. .2289
/organism="Homo sapiens"
/db_xref="Locuid:1827"
/db_xref="taxon:9606"
/clonetype="MGC:10263 IMAGE:394959"
/tissue_type="lung, small cell carcinoma"
/clonelib="NIH MGC 7"
/lab_host="DH10B-R"
/notes="Vector: pOT87"

```

CDS

```

60..1659
/codon_start=1
/product="Down syndrome critical region gene 1"
/protein_id="AA02864.1"
/db_xref="GI:12804023"
/translation="MEEVDLQDLSATIACHLDPRVFDGLCRAPKE
FOYFKPKRVINFSNPFSAADARLQHLKTFEGKGMKLYFACTLIH
KQFLTSPSPSPVGMQVEATPVINDVLLYATLSKLGPGKVELHAA
CESDQKEEEMERMRRPKPIITIRPVTPIHLS"
628 a 465 c 555 g 641 t

```

Query Match 90.4%; Score 2107.4; DB 9; Length 2289;
Best Local Similarity 97.5%; Pred. No. 0;
Matches 2141; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

Qy	135	GAAGCAAGATGCA	TTT	TAGAAAC	TTT	TAAACT	CAGATTTT	TAGCTCCCTGATTTGCCTGTG	194
Db	57	GACTCGAGATGGAGG	TGGAC	CTCGACGAC	CTGCCACGCGCCAC	CACTCGCCTGTAC	116		
Qy	195	GCAAACAGTGTAT	CTTC	CAGCGAAGTGAAC	CAGGGCCAAATTT	TGAGTCCCTCTTTAGG	254		
Db	117	CTGGACCCCGCGGT	TTT	CTGGACGCGCTGTG	CCGGGGCCAAATTT	TGAGTCCCTCTTTAGG	176		
Qy	255	AGTATGACAAAGGAC	ATC	ACCTTTT	TAGAGCTTCAAACGAGT	CAGAAATAAAC	314		
Db	177	AGTATGACAAAGGAC	ATC	ACCTTTT	TAGAGCTTCAAACGAGT	CAGAAATAAAC	236		
Qy	315	TTTCAGCAACCCCTT	CTCG	CAGCAGATGCCAGCTT	CCAGCTGCATAGATGTGATTTCTG	374			
Db	237	TTTCAGCAACCCCTT	CTCG	CAGCAGATGCCAGCTT	CCAGCTGCATAGATGTGATTTCTG	296			
Qy	375	GGAAAGGAATGA	AGTTATATTT	TGCTCAGAC	CTTACACATAGGAAGCTCACACCTGGCT	434			
Db	297	GGAAAGGAATGA	AGTTATATTT	TGCTCAGAC	CTTACACATAGGAAGCTCACACCTGGCT	356			
Qy	435	CCGCCAAATCCAGAC	AAAGAGTTTCT	TGATCTCCCTTCCGCTCT	CCGCCAGTGGGATGG	494			
Db	357	CCGCCAAATCCAGAC	AAAGAGTTTCT	TGATCTCCCTTCCGCTCT	CCGCCAGTGGGATGG	416			
Qy	495	AAACAAGTGGAAAG	ATGCGACCC	CAGTCATAAACT	ATGATCTCTTATATGCCATCTCCAAAG	554			
Db	417	AAACAAGTGGAAAG	ATGCGACCC	CAGTCATAAACT	ATGATCTCTTATATGCCATCTCCAAAG	476			
Qy	585	CTGGGGCCAGGGG	AAAGTATGAAT	TTCAGCAGCAGG	CACTGACCACTCCACGGTGGTG	614			
Db	477	CTGGGGCCAGGGG	AAAGTATGAAT	TTCAGCAGCAGT	CACCACTCCACGGTGGTG	536			

QY	615	GTCCATGTATGTGAGAGTGATCAAGAGAAAGAGAGAAATGGAAGAAATGAGG	674
Db	537	GTCCATGTATGTGAGAGTGATCAAGAGAAAGAGAGAAATGGAAGAAATGAGG	596
QY	675	AGACCTTAGGCCAAAAATTTATCCAGACACAGAGGCCGGAGTACACGCCGATTCACACTCAGC	734
Db	597	AGACCTTAGGCCAAAAATTTATCCAGACACAGAGGCCGGAGTACACGCCGATTCACACTCAGC	656
QY	735	TGAACCTGGCAGCGCAGGAGCGCATTTCCAAATCATACTCACGGGAGGAATCTTTTACTG	794
Db	657	TGAACCTGGCAGCGCAGGAGCGCATTTCCAAATCATACTCACGGGAGGAATCTTTTACTG	716
QY	795	TGGAGGTGGCTGGTTCACGACTCTTTCGGAGGTGGCAGCGGAGATCGGGGTGCAGAAATC	854
Db	717	TGGAGGTGGCTGGTTCACGACTCTTTCGGAGGTGGCAGCGGAGATCGGGGTGCAGAAATC	776
QY	855	CCAGTTCAATGTTGCTCAGAAAGAGAAATCAAGGCCGTCTGCCCTTGTTCATACTGCTGCACAC	914
Db	777	CCAGTTCAATGTTGCTCAGAAAGAGAAATCAAGGCCGTCTGCCCTTGTTCATACTGCTGCACAC	836
QY	915	CAGTTACTGTTTCATGGCACCGGGAATGACTTGGGCCAATCACTGAGTTTGTGGTGATCG	974
Db	837	CAGTTACTGTTTCATGGCACCGGGAATGACTTGGGCCAATCACTGAGTTTGTGGTGATCG	896
QY	975	CACAAGGACATTTGGGACTGTCTTGAGAAAAAGATAAGTAGTGTTTGTGACTTGTTC	1034
Db	897	CACAAGGACATTTGGGACTGTCTTGAGAAAAAGATAAGTAGTGTTTGTGACTTGTTC	956
QY	1035	TTTTCTGTTAGTGTCTGTCTGCCAAGGCGAGTTGATCAGTGAAGCTCAGGAGAGAGCT	1094
Db	957	TTTTCTGTTAGTGTCTGTCTGCCAAGGCGAGTTGATCAGTGAAGCTCAGGAGAGAGCT	1016
QY	1095	TCTCTTTTCTAAGTGGCCTCGACGGGCCACTCTCTACTGGTAGGAAGAGGTACACAGGA	1154
Db	1017	TCTCTTTTCTAAGTGGCCTCGACGGGCCACTCTCTACTGGTAGGAAGAGGTACACAGGA	1076
QY	1155	AGCCGCTAGTGCAGAGAGGTTGTGAAAAACAGCAGCAATGCAATGTGAAATTTGTAGCGT	1214
Db	1077	AGCCGCTAGTGCAGAGAGGTTGTGAAAAACAGCAGCAATGCAATGTGAAATTTGTAGCGT	1136
QY	1215	TTCTCTTTCTCCCTCATGTTCTCATGTTTGTGCATGTATATTCGATTTACAAGACTAA	1274
Db	1137	TTCTCTTTCTCCCTCATGTTCTCATGTTTGTGCATGTATATTCGATTTACAAGACTAA	1196
QY	1275	CCTTGTTCGTATATAAGATTACACGGTTGTTGTTTTACATCTTTTGGGAAGCCAGGAA	1334
Db	1197	CCTTTGTTCGTAATAAAGTTACACGGTTGTTGTTTTACATCTTTTGGGAAGCCAGGAA	1256
QY	1335	GCGTTTGGAAAAAGTATACCTTTCCAGATCTCGGATTTCTCGAATCTTCGACTCTTTGCAACAGCA	1394
Db	1257	GCGTTTGGAAAAAGTATACCTTTCCAGATTTCTCGAATTTCTCGAATCTTCGACTCTTTGCAACAGCA	1316
QY	1395	CTTCTCTGGGAACTCTTCTCGGAATGCAATTCACATCAGCATCCCCACCGTCAACCGTGT	1454
Db	1317	CTTCTCTGGGAACTCTTCTCGGAATGCAATTCACATCAGCATCCCCACCGTCAACCGTGT	1376
QY	1455	AACCTGTGCTTTTGAAAAAGAGTTGATCTGAAATTCCTCTGTAGAAATTTAGCTTATACA	1514
Db	1377	AACCTGTGCTTTTGAAAAAGAGTTGATCTGAAATTCCTCTGTAGAAATTTAGCTTATACA	1436
QY	1515	ATTACAGAAATAGCAGTTTCACTGCGCAACTTTTGTGGGTGAGAAATTTTGTAGTTAGTG	1574
Db	1437	ATTACAGAAATAGCAGTTTCACTGCGCAACTTTTGTGGGTGAGAAATTTTGTAGTTAGTG	1496
QY	1575	TTTGGGATCGGACCTCAGTTCTGTTGTTCTTTTATGTGGTGGTGTCTTATACATGAATC	1634
Db	1497	TTTGGGATCGGACCTCAGTTCTGTTGTTCTTTTATGTGGTGGTGTCTTATACATGAATC	1556
QY	1635	ATAGCCAAAAAATTTTGTGGAAACGTGTGGTGTGAGATAGTTGGTCTTTTACCCACGAA	1694
Db	1557	ATAGCCAAAAAATTTTGTGGAAACGTGTGGTGTGAGATAGTTGGTCTTTTACCCACGAA	1616
QY	1695	GACATCAAGATACACTGTGTAATAAAGCTGATAGCATATATTCATACCTGTCTGACATT	1754

source	1..2284 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="21" /map="21q22.1-q22.2"	
gene	1..2284 /gene="DSCR1" 322..675 /gene="DSCR1" /note="down syndrome candidate region 1; one of four alternatively spliced exon 1"	
CDS		
BASE COUNT	626 a 468 c 534 g 656 t	
ORIGIN		
Query Match	90.3%; Score 2104; DB 9; Length 2284; Best Local Similarity 99.1%; Pred. No. 0; Matches 2116; Conservative 0; Mismatches 20; Indels 0; Gaps 0;	
Qy	196 CAAACAGTATCTTCAGCGAAAGTAAACCGGGCCAAATTTGAGTCCCTTTAGGA	255
Db	134 CAATCTGACATATGGMAAAGCCAAATGTTGGCCAAATTTGATCCCTTTAGGA	193
Qy	256 CGATGCAAGAGCATCACTTCAGTATTTTAAGAGCTTCAACGAGTCAAGATAACT	315
Db	194 CGATGCAAGAGCATCACTTCAGTATTTTAAGAGCTTCAACGAGTCAAGATAACT	253
Qy	316 TCAGCAACCCCTTCTCCGACAGATGCCAGGCTCCAGCTGCAATAAGAGTCTG	375
Db	194 CGATGCAAGAGCATCACTTCAGTATTTTAAGAGCTTCAACGAGTCAAGATAACT	253
Qy	254 TCAGCAACCCCTTCTCCGACAGATGCCAGGCTCCAGCTGCAATAAGAGTCTG	313
Db	194 CGATGCAAGAGCATCACTTCAGTATTTTAAGAGCTTCAACGAGTCAAGATAACT	253
Qy	376 GAAAGGAAATGAAGTTATTTTGCTCAGACCTTACATAGGAAGTCAACCTGCTC	435
Db	314 GAAAGGAAATGAAGTTATTTTGCTCAGACCTTACATAGGAAGTCAACCTGCTC	373
Qy	436 CGCAAAATCCAGACAGAGTTCTGATCTCCCTCCGCTCTCCGCGAGTGGATGA	495
Db	374 CGCAAAATCCAGACAGAGTTCTGATCTCCCTCCGCTCTCCGCGAGTGGATGA	433
Qy	496 AACAAATGGAAGATCCGACCCCATCATTAATCTGATCTCTTATGACCATCTCAAGC	555
Db	434 AACAAATGGAAGATCCGACCCCATCATTAATCTGATCTCTTATGACCATCTCAAGC	493
Qy	556 TGGGGCCAGGGGAAAGTATGAATTTGACGCGAGCACTGACACCACTCCAGCTGCTG	615
Db	494 TGGGGCCAGGGGAAAGTATGAATTTGACGCGAGCACTGACACCACTCCAGCTGCTG	553
Qy	616 TCCATGATGTGAGAGTATCAAGAAAGAGAAAGAAAGAAATGAGGA	675
Db	554 TCCATGATGTGAGAGTATCAAGAAAGAGAAAGAAAGAAATGAGGA	613
Qy	676 GACCTAAGCCAAATTTATCCAGACCGAGAGCCGAGTACACGCCATCCACTCACT	735
Db	614 GACCTAAGCCAAATTTATCCAGACCGAGAGCCGAGTACACGCCATCCACTCACT	673
Qy	736 GAACTGGACGCGAGAGAGCGATTCGAATCATCTCAGGAGGAATCTTTTAACTGT	795
Db	674 GAACTGGACGCGAGAGAGCGATTCGAATCATCTCAGGAGGAATCTTTTAACTGT	733
Qy	796 GAGAGTGGCTGTGTCAGACTTCTTGGAGGTGGAGCCGAGATCGGGGTGGAGAAATCC	855
Db	734 GAGAGTGGCTGTGTCAGACTTCTTGGAGGTGGAGCCGAGATCGGGGTGGAGAAATCC	793
Qy	856 CAGTTATGTGCTCAGAGAGATCAAGCCGTGCTCCCTTTGTTCTAATGCTGACACC	915
Db	794 CAGTTATGTGCTCAGAGAGATCAAGCCGTGCTCCCTTTGTTCTAATGCTGACACC	853
Qy	916 AGTTACTGTTATGCAACCCGGGAATGACTTGGCCCAATCACTGAGTTGTGTGATGCG	975
Db	854 AGTTACTGTTATGCAACCCGGGAATGACTTGGCCCAATCACTGAGTTGTGTGATGCG	913
Qy	976 ACAAGACATTTGGAGCTCTTTGAGAAAACAGATAAGATGATGTTTGTACTGTTCT	1035
Db	914 ACAAGACATTTGGAGCTCTTTGAGAAAACAGATAAGATGATGTTTGTACTGTTCT	973
Qy	1036 TTTCTGTAGTCTCTGTGTCAGAGGAGGAGGAGGAGTGTACGTAGCTCAGAGAGCTT	1095
Db	974 TTTCTGTAGTCTCTGTGTCAGAGGAGGAGGAGGAGTGTACGTAGCTCAGAGAGCTT	1033
Qy	1096 CCGTTTCTAAGTGGCTCTGAGGGCCACTCTCTACTGTTAGGAAGGTACCAAGAA	1155
Db	1034 CCGTTTCTAAGTGGCTCTGAGGGCCACTCTCTACTGTTAGGAAGGTACCAAGAA	1093
Qy	1156 GCGCCATAGGCAAGAGGTTGGAAGACAGCAATGTCGAATTTGAGCGTT	1215
Db	1094 GCGCCATAGGCAAGAGGTTGGAAGACAGCAATGTCGAATTTGAGCGTT	1153
Qy	1216 TCCCTTCTCCCTCATGTTCTCATGTTTGTGATGTAATTAAGTTTCAAGACTAAC	1275
Db	1154 TCCCTTCTCCCTCATGTTCTCATGTTTGTGATGTAATTAAGTTTCAAGACTAAC	1213
Qy	1276 CTTTGTCTGATTAAGTTTACACCGTGTGTTTACATCTTTGGAGACGAGAAAG	1335
Db	1214 CTTTGTCTGATTAAGTTTACACCGTGTGTTTACATCTTTGGAGACGAGAAAG	1273
Qy	1336 CGTTGGAAAACGATACCTTCCAGATTCGGAATTCGATCTTTGCAACAGCAC	1395
Db	1274 CGTTGGAAAACGATACCTTCCAGATTCGGAATTCGATCTTTGCAACAGCAC	1333
Qy	1396 TTGCTTGGGAACCTTCTCGGAATGCAATTCACATCCCAACGTCAGAGTGA	1455
Db	1334 TTGCTTGGGAACCTTCTCGGAATGCAATTCACATCCCAACGTCAGAGTGA	1393
Qy	1456 ACTGTGCTTTTGGCAAAAGATGATCTGMAATTCCTGTAGAAATTTAGCTTAAACA	1515
Db	1394 ACTGTGCTTTTGGCAAAAGATGATCTGMAATTCCTGTAGAAATTTAGCTTAAACA	1453
Qy	1516 TTACAGAAATAGCACTTCTGCAACTTTTATGAGGTGAGAAATTTTATAGGCT	1575
Db	1454 TTACAGAAATAGCACTTCTGCAACTTTTATGAGGTGAGAAATTTTATAGGCT	1513
Qy	1576 TTGGGATCGGACCTCAGTTCTGTTGTTCTTTATGAGGTGAGAAATTTTATAGGCT	1635
Db	1514 TTGGGATCGGACCTCAGTTCTGTTGTTCTTTATGAGGTGAGAAATTTTATAGGCT	1573
Qy	1636 TAGCCAAAACCTTTTGGAAAATGTTGTTGATGATGTTGTTTACCCACGAAG	1695
Db	1574 TAGCCAAAACCTTTTGGAAAATGTTGTTGATGATGTTGTTTACCCACGAAG	1633
Qy	1696 ACATCAAGTATACCTTTGAAATGATGATGATGATGATGATGATGATGATGATGATG	1755
Db	1634 ACATCAAGTATACCTTTGAAATGATGATGATGATGATGATGATGATGATGATGATG	1693
Qy	1756 GGTGAAAGATGAGCAGTGGAGACTAAGATGATTAACCTGTAATCATATGTTG	1815
Db	1694 GGTGAAAGATGAGCAGTGGAGACTAAGATGATTAACCTGTAATCATATGTTG	1753
Qy	1816 TAGAAAAGCTGTTCCATGTTCTAACAGACTTGAATTTAAAGCAATGCAATGATG	1875
Db	1754 TAGAAAAGCTGTTCCATGTTCTAACAGACTTGAATTTAAAGCAATGCAATGATG	1813
Qy	1876 AGATCTGTGCGATATGAGAGGATGAGTGCCTTCCCATTAATCTGATGGAATG	1935
Db	1814 AGATCTGTGCGATATGAGAGGATGAGTGCCTTCCCATTAATCTGATGGAATG	1873
Qy	1936 TTATCTAGGTTAACTTTGTAATTTTCTAGTTGATGATGATGATGATGATGATG	1995
Db	1874 TTATCTAGGTTAACTTTGTAATTTTCTAGTTGATGATGATGATGATGATGATG	1933
Qy	1996 GTATTATTTTGGCTTAATACGTAAATGTTGTCATTTGAAATTAATGTC	2055

Db 1934 GTATTATATTTTGGCCCTTACAATACCGTAACATGTTTGTGTCATTTTGAATAATCTTAATGC 1993
QY 2056 CAAGTAACATGATGCTTTTGGAAATTTGGAAGATGCTTTTATCTTTGAGAGCAATA 2115
Db 1994 CAAGTAACATGATGCTTTTGGAAATTTGGAAGATGCTTTTATCTTTGAGAGCAATA 2053
QY 2116 TGTTTGCATTAATGCTTTTGAATTTGATATCAATCAAGAAATTTGAAGCTTTCTCAAAACC 2175
Db 2054 TGTTTGCATTAATGCTTTTGAATTTGATATCAATCAAGAAATTTGAAGCTTTCTCAAAACC 2113
QY 2176 TGTTTACGCTACTTGTGAAGAGGAGCGGTTTGGAGAGACCATTTGATCGCTGTGCA 2235
Db 2114 TGTTTACGCTACTTGTGAAGAGGAGCGGTTTGGAGAGACCATTTGATCGCTGTGCA 2173
QY 2236 GTGTTTCTTTGTTAAAGTCTTTTAAACTGGAGAGGCTAACTTCAAAATACTTTTAACT 2295
Db 2174 GTGTTTCTTTGTTAAAGTCTTTTAAACTGGAGAGGCTAACTTCAAAATACTTTTAACT 2233
QY 2296 GCATTTCTATAATAAATGGGCACAGTATGCTCTTAC 2331
Db 2234 GCATTTCTATAATAAATGGGCACAGTATGCTCTTAC 2269

RESULT 9
AX410694
LOCUS
DEFINITION Sequence 3341 from Patent WO0229103.
ACCESSION AX410694
VERSION AX410694.1 GI:21443399
KEYWORDS human.
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1
AUTHORS Alvares, C., Horne, D., Peres-da-Silva, S. and Vockley, J. G.
TITLE Gene expression profiles in liver cancer
JOURNAL Patent: WO 0229103-A 3341 11-APR-2002;
GENE LOGIC INC (US)

FEATURES
source
1..2173
/organism="Homo sapiens"
/db_xref="taxon:9606"
/notes="EMBL/GenBank Accession No. U28833"
BASE COUNT 606 a 431 c 503 g 633 t
ORIGIN

Query Match 90.1%; Score 2101; DB 6; Length 2173;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 231 GCCAAATTTGAGTCCCTCTTTAGGAGCTATGACAAGGACATCACCTTTGAGTATTTTAAG 290
Db 58 GCCAAATTTGAGTCCCTCTTTAGGAGCTATGACAAGGACATCACCTTTGAGTATTTTAAG 117
QY 291 AGCTTCAAACGAGTCAAGATAAATTTTCAAGCAACCCCTTCTCCGAGCAGATGCCAGGCTC 350
Db 118 AGCTTCAAACGAGTCAAGATAAATTTTCAAGCAACCCCTTCTCCGAGCAGATGCCAGGCTC 177
QY 351 CAGCTGCATAAGACTGAGTCTTCTGGAAAGGAAATGAAGTATATTTTGTCTCAGACCTTA 410
Db 178 CAGCTGCATAAGACTGAGTCTTCTGGAAAGGAAATGAAGTATATTTTGTCTCAGACCTTA 237
QY 411 CACATAGGAGCTCACACCTGGCTCCGCAAAATCCAGACAGCAGTCTTGTGATCTCCCT 470
Db 238 CACATAGGAGCTCACACCTGGCTCCGCAAAATCCAGACAGCAGTCTTGTGATCTCCCT 297
QY 471 CCGGCTCTCCGAGTGGGATGGAACAAAGTGGAGATGGACCCCAAGTCATAAATCAT 530
Db 298 CCGGCTCTCCGAGTGGGATGGAACAAAGTGGAGATGGACCCCAAGTCATAAATCAT 357
QY 531 GATCTCTATATGCCATCTCCCAAGCTGGGCGAGGGAAGATGAATTCACGAGCG 590

Db 358 GATCTCTTATATGCCATCTCCAAGCTGGGCCAGGGGAAAAGTATGAATTTGACGAGCG 417
QY 591 ACTGACACCACTCCCAAGCTGGTGGTCCATGTATGTGAGAGTATCAAGAGAGGAGAA 650
Db 418 ACTGACACCACTCCCAAGCTGGTGGTCCATGTATGTGAGAGTATCAAGAGAGGAGAA 477
QY 651 GAAGAGAAATGGAAGAATGAGGAGACCTTAAGCCAAAATTTATCCAGACCAAGGAGCGCG 710
Db 478 GAAGAGAAATGGAAGAATGAGGAGACCTTAAGCCAAAATTTATCCAGACCAAGGAGCGCG 537
QY 711 GAGTACACCGCCGATCCACCTCAGCTGAACTGGCA CGCGACGAGGACGCAATTCGAATCAT 770
Db 538 GAGTACACCGCCGATCCACCTCAGCTGAACTGGCA CGCGACGAGGACGCAATTCGAATCAT 597
QY 771 ACTCACGGGAGGAATCTTTTACTGTGGAGTGGCTGGTCAACGACTTCTTCGAGGTGGCA 830
Db 598 ACTCACGGGAGGAATCTTTTACTGTGGAGTGGCTGGTCAACGACTTCTTCGAGGTGGCA 657
QY 831 GCCGAGATCGGGGTGGCAGAAATCCAGTTTCTCAGAGAGAAATCAAGGCCGTG 890
Db 658 GCCGAGATCGGGGTGGCAGAAATCCAGTTTCTCAGAGAGAAATCAAGGCCGTG 717
QY 891 TCCCTTTGTTCTAATGCTGSCACACCAAGTTTACTGTTTCATGGCACCCCGGAAATGACTTGGGC 950
Db 718 TCCCTTTGTTCTAATGCTGSCACACCAAGTTTACTGTTTCATGGCACCCCGGAAATGACTTGGGC 777
QY 951 CAATCAGTGAATTTGTGGTATCGCA CAAGGACATTTGGGACTGTCTTTGAGAAAACAGAT 1010
Db 778 CAATCAGTGAATTTGTGGTATCGCA CAAGGACATTTGGGACTGTCTTTGAGAAAACAGAT 837
QY 1011 AATGATAGTGTGTTTGTACTGTTCTTTTCTGTTAGTGTCTCTGTGCAAGGCGAGTT 1070
Db 838 AATGATAGTGTGTTTGTACTGTTCTTTTCTGTTAGTGTCTCTGTGCAAGGCGAGTT 897
QY 1071 GATCAGTGAATTTGAGGAGAGTCTTCTGTTTCTAAGTGGCTTCGAGGCGCACTCTCTA 1130
Db 898 GATCAGTGAATTTGAGGAGAGTCTTCTGTTTCTAAGTGGCTTCGAGGCGCACTCTCTA 957
QY 1131 CTGTTAGGAAAGGTACCAAGGAGCGCTAGTGAGAGAGTGTGAAAAACAGCAGC 1190
Db 958 CTGTTAGGAAAGGTACCAAGGAGCGCTAGTGAGAGAGTGTGAAAAACAGCAGC 1017
QY 1191 AATGCAATGTGAAATTTGAGCGTTTCTTTTCTCCCTCATGTTCTCATGTTGTGATG 1250
Db 1018 AATGCAATGTGAAATTTGAGCGTTTCTTTTCTCCCTCATGTTCTCATGTTGTGATG 1077
QY 1251 TATATTTACTGATTTTCAAGACTAACTTTGTTGTTATATAAAGTTACACCGTTGTTCTTT 1310
Db 1078 TATATTTACTGATTTTCAAGACTAACTTTGTTGTTATATAAAGTTACACCGTTGTTCTTT 1137
QY 1311 TACATCTTTTGGGAAAGCGGAAAGCGTTTGGAAACGTAATCACTTTCCAGATCTCG 1370
Db 1138 TACATCTTTTGGGAAAGCGGAAAGCGTTTGGAAACGTAATCACTTTCCAGATCTCG 1197
QY 1371 GATTTCTGACTCTTTTGAACAGCAGCTTGTGTTGGGAACTCTTCTCGAATGCAATTCAC 1430
Db 1198 GATTTCTGACTCTTTTGAACAGCAGCTTGTGTTGGGAACTCTTCTCGAATGCAATTCAC 1257
QY 1431 AGCATCCCCAACCGTGAACGTTAACTTTGCTTTGCAAAAGAAAGTTGATCTGAAT 1490
Db 1258 AGCATCCCCAACCGTGAACGTTAACTTTGCTTTGCAAAAGAAAGTTGATCTGAAT 1317
QY 1491 CCTCTGTAGAAATTTAGTGTGTTTGGGATCGGACCTCGGACCTCAGTTCTGTTGTTCTTTA 1550
Db 1318 CCTCTGTAGAAATTTAGTGTGTTTGGGATCGGACCTCGGACCTCAGTTCTGTTGTTCTTTA 1377
QY 1551 GGTGTAGAAATTTAGTGTGTTTGGGATCGGACCTCGGACCTCAGTTCTGTTGTTCTTTA 1610
Db 1378 GGTGTAGAAATTTAGTGTGTTTGGGATCGGACCTCGGACCTCAGTTCTGTTGTTCTTTA 1437
QY 1611 TGTGTGTGTTTCTATACATGAATCATAGCAGGAAATCTTTTGGAAACGTTGTTGTTGAGA 1670
Db 1438 TGTGTGTGTTTCTATACATGAATCATAGCAGGAAATCTTTTGGAAACGTTGTTGTTGAGA 1497

QY 1671 TAGTGGTCTTTTACCCACGACATCAAGATACCTGTAAATTAAGCTGATAGCA 1730
 DB 1498 TAGTGGTCTTTTACCCACGACATCAAGATACCTGTAAATTAAGCTGATAGCA 1557
 QY 1731 TATATTCACTCTGTTGTAACCTTGGGTAAGATAGGAGTGGAGACTAGATAT 1790
 DB 1558 TATATTCACTCTGTTGTAACCTTGGGTAAGATAGGAGTGGAGACTAGATAT 1617
 QY 1791 TAACCACTGTGAATCATATGTGTAGAGAAAGCTGTTCCCATGTCAACAGACTTGA 1850
 DB 1618 TAACCACTGTGAATCATATGTGTAGAGAAAGCTGTTCCCATGTCAACAGACTTGA 1677
 QY 1851 ATTCAGAGATGCAAGTGAATGATGATCTGTGGGATATAGAGGAGTGAAGGCTT 1910
 DB 1678 ATTCAGAGATGCAAGTGAATGATGATCTGTGGGATATAGAGGAGTGAAGGCTT 1737
 QY 1911 TCCCATTTACCTCCGATGGAATGTTATAGTTAAGTTAATTTTCTAGT 1970
 DB 1738 TCCCATTTACCTCCGATGGAATGTTATAGTTAAGTTAATTTTCTAGT 1797
 QY 1971 TGTATGTGATGTCTGTAAATAGTATTAATTTGGCTTACAAATCCGTAACAATG 2030
 DB 1798 TGTATGTGATGTCTGTAAATAGTATTAATTTGGCTTACAAATCCGTAACAATG 1857
 QY 2031 TTTGTCTTTTGAATTACTTAATGCCAAGTAACAAATGCTTTGGAATTTGGAAGAT 2090
 DB 1858 TTTGTCTTTTGAATTACTTAATGCCAAGTAACAAATGCTTTGGAATTTGGAAGAT 1917
 QY 2091 GGTTTATCTTTGAGAAGCAATATGTGTGAATTAATGCTTTGATTTGTCATATCAAG 2150
 DB 1918 GGTTTATCTTTGAGAAGCAATATGTGTGAATTAATGCTTTGATTTGTCATATCAAG 1977
 QY 2151 AATTTGATGAACCTTCTCAAAACCTGTTTAACGCTACTTGTGAAGAGGAGCCGGTTTGG 2210
 DB 1978 AATTTGATGAACCTTCTCAAAACCTGTTTAACGCTACTTGTGAAGAGGAGCCGGTTTGG 2037
 QY 2211 GAGGACCATTTGATCGCTGTCGAAGTGTCTTGTGAAGTCTTTTAACTGAGAGGC 2270
 DB 2038 GAGGACCATTTGATCGCTGTCGAAGTGTCTTGTGAAGTCTTTTAACTGAGAGGC 2097
 QY 2271 TAACCTCAAAATTAATTTTAACTGATCTATTAATTAATGAGGACAGTATGCTCTTA 2330
 DB 2098 TAACCTCAAAATTAATTTTAACTGATCTATTAATTAATGAGGACAGTATGCTCTTA 2157
 QY 2331 C 2331
 DB 2158 C 2158

RESULT 10
 HSU28833 2173 bp mRNA linear PRI 19-APR-2000
 LOCUS Homo sapiens Down syndrome candidate region protein (DSCR1) mRNA,
 DEFINITION complete cds.
 ACCESSION U28833
 VERSION U28833.2 GI:7596916
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens; Chordata; Craniata; Vertebrata; Euteleostomi;
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 2173)
 AUTHORS Puentes,J.J., Pritchard,M.A., Planas,A.M., Bosch,A., Ferrer,I. and
 Estivill,X.
 TITLE A new human gene from the Down syndrome critical region encodes a
 proline-rich protein highly expressed in fetal brain and heart
 JOURNAL Hum. Mol. Genet. 4 (10), 1935-1944 (1995)
 MEDLINE 96121593
 PUBMED 8595418
 REFERENCE 2 (bases 1 to 2173)
 AUTHORS Estivill,X.
 TITLE Direct Submission

JOURNAL Submitted (09-JUN-1995) Xavier Estivill, Molecular Genetics
 Department, Cancer Research Institute, Hospital Duran i Reynals,
 Avia, Castelldefels Km 2.7, Hospitalet, Barcelona, Catalonia 08907,
 Spain
 3 (bases 1 to 2173)
 REFERENCE
 AUTHORS Estivill,X.
 TITLE Direct Submission
 JOURNAL Submitted (19-APR-2000) Xavier Estivill, Molecular Genetics
 Department, Cancer Research Institute, Hospital Duran i Reynals,
 Avia, Castelldefels Km 2.7, Hospitalet, Barcelona, Catalonia 08907,
 Spain
 REMARK Sequence update by submitter
 COMMENT On Apr 19, 2000 this sequence version replaced gi:1125051.
 FEATURES
 source
 location/Qualifiers
 1..2173
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="21"
 /map="21q22.1-q22.2"
 1..2173
 /gene="DSCR1"
 49..564
 /gene="DSCR1"
 /codon_start=1
 /product="down syndrome candidate region protein 1"
 /protein_id="BAB81557.1"
 /db_xref="GI:1125052"
 /translation="MVYARFESLFRDYKDTFQYFKSKRYRINSPFSAADARLQ
 LKHTVEFGKEMKLYFAQLIHGSHLAPRPDQFLISPPAPVGMKQVEDATPVIN
 YDLVLAISKLGPEKYEELHAATDTPSVVVHCSDQEKREBERMERMRPKILOT
 RREYRPIHLS"
 2159
 /gene="DSCR1"
 polyA_site
 BASE COUNT 606 a 431 c 503 g 633 t
 ORIGIN
 Query Match 90.1%; Score 2101; DB 9; Length 2173;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 231 GCCAATTGAGTCCCTCTTTAGACGATGACAAAGACATCACTTCACTATTTTAA 290
 DB 58 GCCAATTGAGTCCCTCTTTAGACGATGACAAAGACATCACTTCACTATTTTAA 117
 QY 291 AGCTTCAACGATGAGATTAATCTTGACGAACCCCTCTCCGACGAGATGCCAGGCTC 350
 DB 118 AGCTTCAACGATGAGATTAATCTTGACGAACCCCTCTCCGACGAGATGCCAGGCTC 177
 QY 351 CAGCTGCATAGACTGAGTTCTGGGAAAGAAATGAATTATTTTGTCAAGCTTA 410
 DB 178 CAGCTGCATAGACTGAGTTCTGGGAAAGAAATGAATTATTTTGTCAAGCTTA 237
 QY 411 CACATGGAAGTCAACCTGCTCCGCCAATCCAGCAAGACTTTCTATCTCCCT 470
 DB 238 CACATGGAAGTCAACCTGCTCCGCCAATCCAGCAAGACTTTCTATCTCCCT 297
 QY 471 CCGGCTCTCCGCACTGGGATGAGAAACAAGTGAAGTCAAGCCCACTATAACTAT 530
 DB 298 CCGGCTCTCCGCACTGGGATGAGAAACAAGTGAAGTCAAGCCCACTATAACTAT 357
 QY 531 GATCTCTTATATGCCATCTTCAAGCTGGGCGCAAGGGAAGATGAATTGACCGCAGG 590
 DB 358 GATCTCTTATATGCCATCTTCAAGCTGGGCGCAAGGGAAGATGAATTGACCGCAGG 417
 QY 591 ACTGACACCACTCCCACTGCTGCTCATATGTAAGTGAAGTGAAGAGAGGAA 650
 DB 418 ACTGACACCACTCCCACTGCTGCTCATATGTAAGTGAAGTGAAGAGAGGAA 477
 QY 651 GAGAGGAATGGAAGATGAGAGACCTTAAGCCAAAATTTCCAGACAGAGAGCGG 710
 DB 478 GAGAGGAATGGAAGATGAGAGACCTTAAGCCAAAATTTCCAGACAGAGAGCGG 537
 QY 711 GAGTACAGCCGATCCACTCACTGTAAGTGGCAACGACAGAGAGCATTTCAATCAT 770

Db 538 GAGTACACCGCGATCACCTCAGCTGAACCTGGCAGCGGACGAGCGCATTCCAAATCAT 597
QY 771 ACTCAGGAGGAATCTTTTACTGAGGAGTGGCTGGTCAACGACTTCTTCGAGGTGGCA 830
Db 598 ACTCAGGAGGAATCTTTTACTGAGGAGTGGCTGGTCAACGACTTCTTCGAGGTGGCA 657
QY 831 GCGGAGATCGGGGTGGCAGAAATCCAGTTCATGTTGCTCAGAGAGAAATCAAGGCCGTG 890
Db 658 GCGGAGATCGGGGTGGCAGAAATCCAGTTCATGTTGCTCAGAGAGAAATCAAGGCCGTG 717
QY 891 TCCCTCTGTTCTAAATGCTGCACACAGTACTTCTTCTGAGTGGCTGCTTGGAAACACAGAT 1010
Db 778 CAATCACTGAGTTGTGGTGATCGCACAGGACATTTGGGACTGCTTGGAGAAACAGAT 837
QY 1011 AATGATAGTGTGTTGACTTCTTCTTCTGAGTGGTCTGCTGCGCAAGGCGAGTT 1070
Db 838 AATGATAGTGTGTTGACTTCTTCTTCTGAGTGGTCTGCTGCGCAAGGCGAGTT 897
QY 1071 GATCAGTGAAGTCAAGGAGAGTCTTCTGTTTCTAAGTGGCTGCGAGGCGCACTCTCTA 1130
Db 898 GATCAGTGAAGTCAAGGAGAGTCTTCTGTTTCTAAGTGGCTGCGAGGCGCACTCTCTA 957
QY 1131 CTGGTAGGAAGAGTACACAGGAAGCGGCTAGTCAGAGAGAGTGTGAAACAGCAGC 1190
Db 958 CTGGTAGGAAGAGTACACAGGAAGCGGCTAGTCAGAGAGAGTGTGAAACAGCAGC 1017
QY 1191 AATGCAATGTGAAATGTAGCGTTTCTTCTTCTTCTGAGTGGTCTGCTGCGCAAGGCGAGTT 1250
Db 1018 AATGCAATGTGAAATGTAGCGTTTCTTCTTCTTCTGAGTGGTCTGCTGCGCAAGGCGAGTT 1077
QY 1251 TATATTACTGATTACAGACTAACTTTGTTGTTATATAAAGTTACACCGTTGTTGTTT 1310
Db 1078 TATATTACTGATTACAGACTAACTTTGTTGTTATATAAAGTTACACCGTTGTTGTTT 1137
QY 1311 TACATCTTTTGGGAAGCCAGGAAGCGTTTGGAAAGATATACCTTTCCAGATTCTCG 1370
Db 1138 TACATCTTTTGGGAAGCCAGGAAGCGTTTGGAAAGATATACCTTTCCAGATTCTCG 1197
QY 1371 GATTCTCGACTCTTTGCAACAGCAGTCTGCTTGGCGAATCTTCTCGAATGCAATTCACATC 1430
Db 1198 GATTCTCGACTCTTTGCAACAGCAGTCTGCTTGGCGAATCTTCTCGAATGCAATTCACATC 1257
QY 1431 AGCATCCCCAACCGTGCACAGTGTAACTTGTGCTTTTGCAAAAGAAAGTTGATCTGAAAT 1490
Db 1258 AGCATCCCCAACCGTGCACAGTGTAACTTGTGCTTTTGCAAAAGAAAGTTGATCTGAAAT 1317
QY 1491 CCTCTGTAGAAATTTAGCTTATACAAATTCAGAGATAGCAGTTTCACTGCGCAACTTTTAGT 1550
Db 1318 CCTCTGTAGAAATTTAGCTTATACAAATTCAGAGATAGCAGTTTCACTGCGCAACTTTTAGT 1377
QY 1551 GGGTGAGAAATTTAGTTAGTGTGTTGGGATCGGACCTCAGTTTCTGTTGTTTCTTTT 1610
Db 1378 GGGTGAGAAATTTAGTTAGTGTGTTGGGATCGGACCTCAGTTTCTGTTGTTTCTTTT 1437
QY 1611 TGTGTTGTTTCTATACATGATCATAGCCAAAACCTTTTGGAAACCTGTTGGTGGAGA 1670
Db 1438 TGTGTTGTTTCTATACATGATCATAGCCAAAACCTTTTGGAAACCTGTTGGTGGAGA 1497
QY 1671 TAGTTGGTTCTTTTACCCACAGACATCAAGATATACCTTCTTAATAAGCTGATAGCA 1730
Db 1498 TAGTTGGTTCTTTTACCCACAGACATCAAGATATACCTTCTTAATAAGCTGATAGCA 1557
QY 1731 TATATTCACTCTGTTGTTACCTTGGTGAAAGATATGSCAGTGGGAGACTAAGATGTAT 1790
Db 1558 TATATTCACTCTGTTGTTACCTTGGTGAAAGATATGSCAGTGGGAGACTAAGATGTAT 1617
QY 1791 TAACTTACCTGTGAAATCATATGCTGTTAGGAAGAGTGTCCCATGCTTAAACAGGACTTGA 1850

Db 1618 TAACCTACCTGTGAATCATATGTTGTAGGAAAAAGCTGTTCCCATGCTTAACAGGACTTGA 1677
QY 1851 ATTCAAAGCATGTCAAAGTGATAGTATCTGTGGCGATATGAGAGGATGCAAGTGCCTT 1910
Db 1678 ATTCAAAGCATGTCAAAGTGATAGTATCTGTGGCGATATGAGAGGATGCAAGTGCCTT 1737
QY 1911 TCCCATTCATTCCTCATGAGAAATTTTACTAGGTTAAACATTTGTTTCTAGT 1970
Db 1738 TCCCATTCATTCCTCATGAGAAATTTTACTAGGTTAAACATTTGTTTCTAGT 1797
QY 1971 TGTAAATGTGTATGCTGTGTAATAGGTATATATATTTTGGCCTTACAATACCGTACAATG 2030
Db 1798 TGTAAATGTGTATGCTGTGTAATAGGTATATATTTTGGCCTTACAATACCGTACAATG 1857
QY 2031 TTTGTCAATTTTGAATACCTTAATGCAAGTAAACAATGCAATGCTTTGGAAATTTGGAAGAT 2090
Db 1858 TTTGTCAATTTTGAATACCTTAATGCAAGTAAACAATGCAATGCTTTGGAAATTTGGAAGAT 1917
QY 2091 GGTTTTATTTCTTTGAGAAGCAATATGTTTGCATTTAAATGCTTTGATTTTCATATCAAG 2150
Db 1918 GGTTTTATTTCTTTGAGAAGCAATATGTTTGCATTTAAATGCTTTGATTTTCATATCAAG 1977
QY 2151 AAATGATTGAACGTTCTCAAAACCTGTTTACGTTACTTTGTAAGAGGAGCGGTTTGG 2210
Db 1978 AAATGATTGAACGTTCTCAAAACCTGTTTACGTTACTTTGTAAGAGGAGCGGTTTGG 2037
QY 2211 GAGAGACCATTCATCGCTGCTCAAGTGTCTTCTTAAAGTGTCTTTAACTTGGAGAGGC 2270
Db 2038 GAGAGACCATTCATCGCTGCTCAAGTGTCTTCTTAAAGTGTCTTTAACTTGGAGAGGC 2097
QY 2271 TAACTTCAAAATACATTTTAACTGCAATTCATATAAATGGAAGGACAGTATGCTCCTTA 2330
Db 2098 TAACTTCAAAATACATTTTAACTGCAATTCATATAAATGGAAGGACAGTATGCTCCTTA 2157
QY 2331 C 2331
Db 2158 C 2158

AK055845 2679 bp mRNA linear PRI 01-AUG-2002
RESULT 11 Homo sapiens CDNA FLJ31283 fis, clone KIDNE2006820, highly similar
AK055845 LOCUS TO DOWN SYNDROME CRITICAL REGION PROTEIN 1.
DEFINITION
AK055845
ACCESSION
VERSION
KEYWORDS
SOURCE
Oligo capping; fis (full insert sequence).
Homo sapiens kidney cDNA to mRNA, clone_lib:KIDNE2
Clone:KIDNE2006820.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS
Tashiro,H., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S.,
Fukuzumi,Y., Fujimori,Y., Komiyama,M., Sugiyama,T., Irie,R.,
Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y.,
Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H.,
Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K.,
Wagatsuna,M., Murakawa,K., Kanenori,K., Takahashi-Fujii,A.,
Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S.,
Nagahara,K., Masuho,Y., Nagai,K. and Isogai,T.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 2679)
Isogai,T., Otsuki,T. and Sugiyama,T.
Direct Submission
Submitted (24-Oct-2001) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan

Key Technology Center etc.); 5'- & 3'-end one pass sequencing; RAB, HRI, and Biotechnology Center; National Institute of Technology and Evaluation; Clone selection for full insert sequencing; RAB and HRI.

FEATURES

source

Location/Qualifiers -

1. 2679
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="KIDNE200620"
/issue_type="Kidney"
/clone_lib="KIDNE2"
/note="Cloning vector: PME18SFL3"

BASE COUNT 739 a 540 c 631 g 769 t
ORIGIN

Query Match 89.8%; Score 2093.6; DB 9; Length 2679;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 2096; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 230 GGCCTAATTGAGTCCCTCTTTAGGACGTATGACAGGACATCCTTCAGATTTTAA 289
DB |||||
QY 580 GGCCTAATTGAGTCCCTCTTTAGGACGTATGACAGGACATCCTTCAGATTTTAA 639
DB |||||
QY 290 GAGCTTCAACGAGTCAGATTAACCTTCAGCAACCCCTTCTCCGACAGATGCCAGCT 349
DB 640 GAGCTTCAACGAGTCAGATTAACCTTCAGCAACCCCTTCTCCGACAGATGCCAGCT 699
QY 350 CCAGCTGCATAGACTGAGTTTCTGGGAAAGGAAATGAAATTATTTTGCTCAGACCTT 409
DB 700 CCAGCTGCATAGACTGAGTTTCTGGGAAAGGAAATGAAATTATTTTGCTCAGACCTT 759
QY 410 ACACATAGGAGGTCACACCTGGCTCCGCAATCCAGACAGACAGTTCTGATCTCCC 469
DB 760 ACACATAGGAGGTCACACCTGGCTCCGCAATCCAGACAGACAGTTCTGATCTCCC 819
QY 470 TCCCGCTCTCCGCAAGTGGATGGAAACAAATGGAAGATGCCACCTTAACCTA 529
DB 820 TCCCGCTCTCCGCAAGTGGATGGAAACAAATGGAAGATGCCACCTTAACCTA 879
QY 530 TGAATCTTTATATGCAATCTTCGAAGCTGGGGCCAGGGGAAAGTATGATTCACAGCAGC 589
DB 880 TGAATCTTTATATGCAATCTTCGAAGCTGGGGCCAGGGGAAAGTATGATTCACAGCAGC 939
QY 590 GACTGACACCACTCCCAAGCTGTGTGTCATGTATGTGAGATGATCAGAGAGAGGA 649
DB 940 GACTGACACCACTCCCAAGCTGTGTGTCATGTATGTGAGATGATCAGAGAGAGGA 999
QY 650 AGAAGAGAAATGGAAGAAATGAGAGACCTTAACCAAAATTTATCCAGACCCAGAGAGCC 709
DB 1000 AGAAGAGAAATGGAAGAAATGAGAGACCTTAACCAAAATTTATCCAGACCCAGAGAGCC 1059
QY 710 GGAATACACGCGCGATCCACTCAGCTGAATCGACCGGACGAGAGAGCATTTCCAAATCA 769
DB 1060 GGAATACACGCGCGATCCACTCAGCTGAATCGACCGGACGAGAGAGCATTTCCAAATCA 1119
QY 770 TACTCAGCGGAGGAATCTTTTACTGTGAGGTGTGCTGTCAGACTTCTTGGAGGTGGC 829
DB 1120 TACTCAGCGGAGGAATCTTTTACTGTGAGGTGTGCTGTCAGACTTCTTGGAGGTGGC 1179
QY 830 AGCCGAGATCGGGGTGGGAGAAATCCAGTTTCTGCTCAGAGAGAAATCAAGGCGCT 889
DB 1180 AGCCGAGATCGGGGTGGGAGAAATCCAGTTTCTGCTCAGAGAGAAATCAAGGCGCT 1239
QY 890 GTCCCTGTTCTTAATGCTGACACCACTTACTGTTTCAATGACCCCGGAAATGACTTGGG 949
DB 1240 GTCCCTGTTCTTAATGCTGACACCACTTACTGTTTCAATGACCCCGGAAATGACTTGGG 1299
QY 950 CCAATCAGTACTTGTGTGTATGTCACAGAGACATTTGGGACTGTCTTGGAGAAACAGA 1009
DB 1300 CCAATCAGTACTTGTGTGTATGTCACAGAGACATTTGGGACTGTCTTGGAGAAACAGA 1359
QY 1010 TAAATAGATGTTTGTACTTGTCTTTCTGTGAGTTCTGTGTCTGTCAAGGAGCAGT 1069

DB 1360 TAAATAGATGTTTGTACTTGTCTTTCTGTGAGTTCTGTGTCTGTCAAGGAGCAGT 1419
QY 1070 TGAATCAGTACTCAGAGAGAGCTTCTGTGTTCTTAAGTGGCTGAGAGGACACTCTCT 1129
DB 1420 TGAATCAGTACTCAGAGAGAGCTTCTGTGTTCTTAAGTGGCTGAGAGGACACTCTCT 1479
QY 1130 ACTGTAGAGAGAGTACACAGAGAACCCGCTTATGTCAGAGAGGTTTGAAACAGCAG 1189
DB 1480 ACTGTAGAGAGAGTACACAGAGAACCCGCTTATGTCAGAGAGGTTTGAAACAGCAG 1539
QY 1190 CAATGCAATGGAAGAAATGAGAGGTTCTTCTTCCCATGTTCTCATGTTTGTGCAT 1249
DB 1540 CAATGCAATGGAAGAAATGAGAGGTTCTTCTTCCCATGTTCTCATGTTTGTGCAT 1599
QY 1250 GTATATTACTGATTTTACAGACTAACCTTTGTTGATATAAAGTTACACCGTTGTTGTT 1309
DB 1600 GTATATTACTGATTTTACAGACTAACCTTTGTTGATATAAAGTTACACCGTTGTTGTT 1659
QY 1310 TTACATCTTTTGGGAGACCCAGAGAAAGCTTTGGAGAAACGATACCTTCCGATTTCTC 1369
DB 1660 TTACATCTTTTGGGAGACCCAGAGAAAGCTTTGGAGAAACGATACCTTCCGATTTCTC 1719
QY 1370 GGATTTCTGACTTTTTCACACAGACTTGGTGGCAACTCTTCTGGAATGCAATTCAT 1429
DB 1720 GGATTTCTGACTTTTTCACACAGACTTGGTGGCAACTCTTCTGGAATGCAATTCAT 1779
QY 1430 CAGATCCCAACCGTGCACAGCTGTAACTTGTCTTTGCAAAAGAAAGTTGATCTGAAT 1489
DB 1780 CAGATCCCAACCGTGCACAGCTGTAACTTGTCTTTGCAAAAGAAAGTTGATCTGAAT 1839
QY 1490 TCCCTGTGAGAAATTTAGGTTTATACATTCAGAGAAATGACAGTTTCACTGCCAATTTAG 1549
DB 1840 TCCCTGTGAGAAATTTAGGTTTATACATTCAGAGAAATGACAGTTTCACTGCCAATTTAG 1899
QY 1550 TGGGTGAGAAATTTTATGTTAGTGTGGATCGAAGCTTCAGTTCTGTGTTCTTTT 1609
DB 1900 TGGGTGAGAAATTTTATGTTAGTGTGGATCGAAGCTTCAGTTCTGTGTTCTTTT 1959
QY 1610 ATGTGTGTTTCTATATACATGATCATGACCAAAACCTTTTGGAAACGTGTGGAG 1669
DB 1960 ATGTGTGTTTCTATATACATGATCATGACCAAAACCTTTTGGAAACGTGTGGAG 2019
QY 1670 ATAGTTGTTCTTTTACCAGCAAGAGATCAGATACACTTGTAAATTAAGCTGATAGC 1729
DB 2020 ATAGTTGTTCTTTTACCAGCAAGAGATCAGATACACTTGTAAATTAAGCTGATAGC 2079
QY 1730 ATATATCTATACCTGTGTGATCACTTGGGTGAAGATATGCGAGTGGAGACTTAAGATTA 1789
DB 2080 ATATATCTATACCTGTGTGATCACTTGGGTGAAGATATGCGAGTGGAGACTTAAGATTA 2139
QY 1790 TTAACCTATCTGTGATATCAATGTTGTAGAAACCTGTTCCATGTCTTAACAGAGACTTG 1849
DB 2140 TTAACCTATCTGTGATATCAATGTTGTAGAAACCTGTTCCATGTCTTAACAGAGACTTG 2199
QY 1850 AATTCAACCATGTCAAGTGTATGATCTGTGGCATATGAGAGGATGCAAGTGCCT 1909
DB 2200 AATTCAACCATGTCAAGTGTATGATCTGTGGCATATGAGAGGATGCAAGTGCCT 2259
QY 1910 TTCCCATTCATCTCTGATGGAATGTTTACTAGTTTAACTTTGTTTCTAG 1969
DB 2260 TTCCCATTCATCTCTGATGGAATGTTTACTAGTTTAACTTTGTTTCTAG 2319
QY 1970 TTGTATATGATATGCTGTGTAATAGATTAATTTTGGCTTCAATACCGTAACAT 2029
DB 2320 TTGTATATGATATGCTGTGTAATAGATTAATTTTGGCTTCAATACCGTAACAT 2379
QY 2030 GTTGTGATTTTGAATTAATCTTAATGCAAGTAACATGCTTTTGGAAATTTGGAATA 2089
DB 2380 GTTGTGATTTTGAATTAATCTTAATGCAAGTAACATGCTTTTGGAAATTTGGAATA 2439
QY 2090 TGGTTTATCTTTTGAAGAGCAATATGTTTGAATTAATGCTTTGATTTGATATCA 2149
DB 2440 TGGTTTATCTTTTGAAGAGCAATATGTTTGAATTAATGCTTTGATTTGATATCA 2499

QY 2150 GAAATTGATTGAACGTTCTCAAAACCTGTTTACGGTACTTGGTAAGAGGAGCGGTTTG 2209
Db 2500 GAAATTGATTGAACGTTCTCAAAACCTGTTTACGGTACTTGGTAAGAGGAGCGGTTTG 2559
QY 2210 GGAGAGACCAATTGCAATCGCTGCTCCAAAGTGTCTTCTGTTAAAGTCTTTTAAACTGGAGAGG 2269
Db 2560 GGAGAGACCAATTGCAATCGCTGCTCCAAAGTGTCTTCTGTTAAAGTCTTTTAAACTGGAGAGG 2619
QY 2270 CTAACCTCAAAATACATTTTTTTTAACTGCAATCTATAATAAATGGGCACAGTATGCTCCTT 2329
Db 2620 CTAACCTCAAAATATTTTTTTTAACTGCAATCTATAATAAATGGGCACAGTATGCTCCTT 2679

RESULT 12
AR034241
LOCUS AR034241 2174 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 1 from patent US 5869318.
ACCESSION AR034241
VERSION AR034241.1 GI:5949846
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 2174)
Palleja,X.Estivill., Fuentes,J.Jose. and Pritchard,M.
Gene sequence of the Down syndrome critical region of human
chromosome 21, identified by a new 'Alu-splicing PCR' technique,
coding for a proline-rich protein (DSCR1) highly expressed in
foetal brain and in heart and method for characterizing it
Patent: US 5869318-A 1 09-FEB-1999;
Location/Qualifiers
1..2174
/organism="unknown"

BASE COUNT 605 a 436 c 502 g 628 t 3 others

Query Match 88.3%; Score 2059.2; DB 6; Length 2174;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 2081; Conservative 0; Mismatches 20; Indels 1; Gaps 1;

QY 231 GCCAATTTGAGTCCCTTTAGAGCGTATGCAAGGACATCACCTTTTCAGTATTTTAAG 290
Db 58 GCCAATTTGAGTCCCTTTAGAGCGTATGCAAGGACATCACCTTTTCAGTATTTTAAG 117
QY 291 AGCTTCAAACGAGTCAGATAAATCTTCAAGCAACCCCTTCTCCGACGACATGCCAGGCTC 350
Db 118 AGCTTCAAACGAGTCAGATAAATCTTCAAGCAACCCCTTCTCCGACGACATGCCAGGCTC 177
QY 351 CAGCTGATAAAGCTGAGTTTCTGGAAAGGAAATGAAGTTATATTTTCTCAGACCTTA 410
Db 178 CAGCTGATAAAGCTGAGTTTCTGGAAAGGAAATGAAGTTATATTTTCTCAGACCTTA 237
QY 411 CACATAGGAAGCTCACACCTGGCTCCGCCAATTCAGACAGAGTTCCTGATCTCCCT 470
Db 238 CACATAGGAAGCTCACACCTGGCTCCGCCAATTCAGACAGAGTTCCTGATCTCCCT 297
QY 471 CCGCCTCTCCGCACTGGGATGGAAACAACTGGAAGATGCCACCCAGTCATAACTAT 530
Db 298 CCGCCTCTCCGCACTGGGATGGAAACAACTGGAAGATGCCACCCAGTCATAACTAT 357
QY 531 GATCTCTTATATGCCATCTCAAGCTGGGGCCAGGGGAAAGTATGAATTCAGCGACGC 590
Db 358 GATCTCTTATATGCCATCTCAAGCTGGGGCCAGGGGAAAGTATGAATTCAGCGACGC 417
QY 591 ACTGACACCACTCCACAGCTGTGTGCTCAATGATGTGAGAGTATCAAGAGAGGAA 650
Db 418 ACTGACACCACTCCACAGCTGTGTGCTCAATGATGTGAGAGTATCAAGAGAGGAA 477
QY 651 GAAGAGGAATGGAAAGATAGGAGACCTTAAGCCAAAATTTTCCAGACCCAGGAGCCG 710
Db 478 GAAGAGGAATGGAAAGATAGGAGACCTTAAGCCAAAATTTTCCAGACCCAGGAGCCG 537

QY 711 GAGTACACGCGCGATCCACCTCAGCTGAACCTGCAACGCGAGGAGCGCATTCACAAATCAT 770
Db 538 GAGTACACGCGCGATCCACCTCAGCTGAACCTGCAACGCGAGGAGCGCATTCACAAATCAT 597
QY 771 ACTCACGGGAGGAATCTTTTACTGTGGAGGTGGCTGGTCAAGACTTCTTCGGAGGTGGCA 830
Db 598 ACTCACGGGAGGAATCTTTTACTGTGGAGGTGGCTGGTCAAGACTTCTTCGGAGGTGGCA 657
QY 831 GCCGAGATCGGGGTGGCAGAAATCCAGATTCAATGTTGCTCAGAAGAGAAATCAAGGCCGTG 890
Db 658 GCCGAGATCGGGGTGGCAGAAATCCAGATTCAATGTTGCTCAGAAGAGAAATCAAGGCCGTG 717
QY 891 TCCCTCTGTTCTAAATGCTGCGACACCAAGTTACTGTTCAATGGCACCCGGGAATCACTTGGGC 950
Db 718 TCCCTCTGTTCTAAATGCTGCGACACCAAGTTACTGTTCAATGGCACCCGGGAATCACTTGGGC 777
QY 951 CAATCACTGAGTTTGTGGTGTGATCGCACAAGGACATTTTGGGACTGTCTTGAGAAAACAGAT 1010
Db 778 CAATCACTGAGTTTGTGGTGTGATCGCACAAGGACATTTTGGGACTGTCTTGAGAAAACAGAT 837
QY 1011 AATGATAGTGTGTTGTACTGTTCTTTTCTGGTAGGTTCTGTTGCTGCAAGGGCAGGTT 1070
Db 838 AATGATAGTGTGTTGTACTGTTCTTTTCTGGTAGGTTCTGTTGCTGCAAGGGCAGGTT 897
QY 1071 GATCAGTGAGCTCAGGAGAGAGCTTCTGTTTCTTAAGTGGCTGCGAGGGCCACTCTCTTA 1130
Db 898 GATCAGTGAGCTCAGGAGAGAGCTTCTGTTTCTTAAGTGGCTGCGAGGGCCACTCTCTTA 957
QY 1131 CTGGTAGGAAGAGGTGACCAAGGAGCGCTAGTGAGAGAGGTTGTGAAAACAGGAGC 1190
Db 958 CTGGTAGGAAGAGGTGACCAAGGAGCGCTAGTGAGAGAGGTTGTGAAAACAGGAGC 1017
QY 1191 AATGCAATGTCGAAATGTAGCGTTTCTTTCTCCCTCATGTTCTCATGTTTGTGTCATG 1250
Db 1018 AATGCAATGTCGAAATGTAGCGTTTCTTTCTCCCTCATGTTCTCATGTTTGTGTCATG 1077
QY 1251 TATATTACTGATTTTACAAGACTAACCTTTGTTTGGTATATAAAGTTACACCGTTGTTGTTT 1310
Db 1078 TATATTACTGATTTTACAAGACTAACCTTTGTTTGGTATATAAAGTTACACCGTTGTTGTTT 1137
QY 1311 TACATCTTTTGGGAAGCGAGAAAGCGTTTGGAAACGATATCACCTTTTCCAGATTTCTCG 1370
Db 1138 TACATCTTTTGGGAAGCGAGAAAGCGTTTGGAAACGATATCACCTTTTCCAGATTTCTCG 1197
QY 1371 GATTCTCGACTCTTTGCAACAGCACTTGTGCGGAACTCTTCTGGAATGCAATTCACCT 1430
Db 1198 GATTCTCGACTCTTTGCAACAGCACTTGTGCGGAACTCTTCTGGAATGCAATTCACCT 1257
QY 1431 AGCATCCCCAACCGTGCACGTTGTAACCTTGTGCTTTTGGCAAAAGAGTTGATCTGAAATT 1490
Db 1258 AGCATCCCCAACCGTGCACGTTGTAACCTTGTGCTTTTGGCAAAAGAGTTGATCTGAAATT 1317
QY 1491 CCTCTGTAGAAATTTAGCTTATACAATTCAGAGAAATAGCAGTTTCACTGCCAACTTTTAGT 1550
Db 1318 CCTCTGTAGAAATTTAGCTTATACAATTCAGAGAAATAGCAGTTTCACTGCCAACTTTTAGT 1377
QY 1551 GGGTGGAAATTTTAGTTTGGATTCGGACCTCGGACCTCAGTTTCTGTTGTTCTTTTA 1610
Db 1378 GGGTGGAAATTTTAGTTTGGATTCGGATTCGGACCTCAGTTTCTGTTGTTCTTTTA 1437
QY 1611 TGTGGTGGTTTCTATACATGAATCATAGCCAAAACCTTTTGGAAACCTGTTGTTGTTGAGA 1670
Db 1438 TGTGGTGGTTTCTATACATGAATCATAGCCAAAACCTTTTGGAAACCTGTTGTTGTTGAGA 1497
QY 1671 TAGTGTGTTTCTTTTACCACCAAGACATCAAGATACACTTGTAAATAAGCTGATGAGA 1730
Db 1498 TAGTGTGTTTCTTTTACCACCAAGACATCAAGATACACTTGTAAATAAGCTGATGAGA 1557
QY 1731 TATATTCTACCTGTTGTACACTTGGGTGAAAGTATGGCAGTGGGAGACTAAGATGATAT 1790
Db 1558 TATATTCTACCTGTTGTACACTTGGGTGAAAGTATGGCAGTGGGAGACTAAGATGATAT 1617
QY 1791 TAACTTACCTGTGTAATCATATGTTGTAGGAAAGCTGTTCCCATGCTTAAACAGGACTTGA 1850

Db 1618 TAACCTACCTGGAATCATATGTTAGAGAAAAGTGTCCCAATCTCAACAGACTTGA 1677
 QY 1851 ATTCAACAGATGTCAGTATGATAGATCTGTGCGATATGAGGGATGCAGTCCCTT 1910
 Db 1678 ATTCAACAGATGTCAGTATGATAGATCTGTGCGATATGAGGGATGCAGTCCCTT 1737
 QY 1911 TCCCATTCATTCCTGATGAGAAATGTTATACAGTAAACAATTTGTAATTTTCTAGT 1970
 Db 1738 TCCGATTCATTCCTGATGAGAAATGTTATACAGTAAACAATTTGTAATTTTCTAGT 1797
 QY 1971 TGTATGTCATGCTGTAATAGGATATATTTGGCTTACAAATACCGTAAACAATG 2030
 Db 1798 TGTATGTCATGCTGTAATAGGATATATTTGGCTTACAAATACCGTAAACAATG 1857
 QY 2031 TTTGTCATTTTGAATACTTAAATGCCAGTAAACAATGCTTTGGAAATTTGGAAAT 2090
 Db 1858 TTTGTCATTTTGAATACTTAAATGCCAGTAAACAATGCTTTGGAAATTTGGAAAT 1917
 QY 2091 GGTATATCTTCTTGAAGCAAAATATGTTGCAATTAATGCTTTGATGTTGATCAATG 2150
 Db 1918 GGTATATCTTCTTGAAGCAAAATATGTTGCAATTAATGCTTTGATGTTGATCAATG 1977
 QY 2151 AATGATGTAAGAGCTTCTCAACCTGTTACCGTACTTGTAAAGAGGAGCGGTTTGG 2210
 Db 1978 AATGATGTAAGAGCTTCTCAACCTGTTACCGTACTTGTAAAGAGGAGCGGTTTGG 2037
 QY 2211 GAGAGACCATTCGATGCTT-GTCCAAAGTCTTCTTGAAGTCTTTTAAACCTGAGAGG 2269
 Db 2038 GAGAGACCATTCGATGCTT-GTCCAAAGTCTTCTTGAAGTCTTTTAAACCTGAGAGG 2097
 QY 2270 CTAAACCTCAAAATCTTTTAACTGATCTTATTAATTAATGAGGACAGTATGCTCTT 2329
 Db 2098 CTAAACCTCAAAATCTTTTAACTGATCTTATTAATTAATGAGGACAGTATGCTCTT 2157
 QY 2330 AC 2331
 Db 2158 AC 2159

RESULT 13
 AY007090/c 1792 bp mRNA linear PRI 31-AUG-2000
 LOCUS Homo sapiens clone TCCCIA00112 mRNA sequence.
 DEFINITION AY007090
 ACCESSION AY007090.1 GI:9955979
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

Homo sapiens.
 Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS 1 (bases 1 to 1792)
 TITLE Anderson, B., Wentland, M.A., Rifeferente, J.Y., Liu, W., and Gibbs, R.A.
 JOURNAL Anal. Biochem. 236 (1), 107-113 (1996)
 MEDLINE 96207227
 PUBMED 8619474

REFERENCE
 AUTHORS 2 (bases 1 to 1792)
 TITLE Yu, W., Anderson, B., Morley, K.C., Muzny, D.M., Ding, Y., Liu, W.,
 JOURNAL Rifeferente, J.Y., Wentland, M.A., Lennon, G., and Gibbs, R.A.
 MEDLINE Large-scale concatenation cDNA sequencing
 PUBMED 97264341
 9110174

REFERENCE
 AUTHORS 3 (bases 1 to 1792)
 TITLE Zhou, J., Yu, W., Tang, H., Mei, G., Tsang, Y.T.M., Bouck, J., Gibbs, R.A.
 JOURNAL Direct Submision
 77030, USA

COMMENT The clone request should be directed to Dr. J. Margolin at
 Pediatrics-Hematology & Oncology, Texas Children's Center

102514, Houston, Texas 77030, USA. Telephone: 713-770-4583
 margolin@bcm.tmc.edu.
 Location/Qualifiers
 source
 1. 1792
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="TCCCIA00112"
 /clone_11b="constructed by Y.T.M. Tsang"
 /dev_stage="infant"
 /note="from patient with acute lymphoblastic leukemia"
 BASE COUNT 547 a 428 c 335 g 482 t
 ORIGIN
 Query Match 76.8%; Score 1790.4; DB 9; Length 1792;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1791; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 536 CTTATATGATCTCCCAAGCTGGGGCCAGGAAAGTATGATATTCACGACGACTGA 595
 Db 1792 CTTATATGATCTCCCAAGCTGGGGCCAGGAAAGTATGATATTCACGACGACTGA 1733
 QY 596 CACCACTCCGACGCTGTGTGTCATGTATGTAGATGATCAAGAGAGAGAGAGAGAG 655
 Db 1732 CACCACTCCGACGCTGTGTGTCATGTATGTAGATGATCAAGAGAGAGAGAGAGAG 1673
 QY 656 GGAATGGAAGAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 715
 Db 1672 GGAATGGAAGAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1613
 QY 716 CACGCGATCCACCTCAGCTGAGTAACTGACGACGACGACGACGACGACGACGAC 775
 Db 1612 CACGCGATCCACCTCAGCTGAGTAACTGAGTAACTGAGTAACTGAGTAACTG 1553
 QY 776 CCGAGAGAACTTTTACTGTGTGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 835
 Db 1552 CCGAGAGAACTTTTACTGTGTGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1493
 QY 836 GATCGGGGTGGCAGAAATCCCAAGTATGTTGTCAGAGAGAAATCAAGCCGTGTC 895
 Db 1492 GATCGGGGTGGCAGAAATCCCAAGTATGTTGTCAGAGAGAAATCAAGCCGTGTC 1433
 QY 896 TTGTTCTAATGCTGACACACAGTTACTGTCATGAGCACCAGGAAATGACTGGCCAT 955
 Db 1432 TTGTTCTAATGCTGACACACAGTTACTGTCATGAGCACCAGGAAATGACTGGCCAT 1373
 QY 956 ACTGAGTTGTGTGTATGAGCACAAGAGATTTGGGACTGCTTGAGAAACAGATATGA 1015
 Db 1372 ACTGAGTTGTGTGTATGAGCACAAGAGATTTGGGACTGCTTGAGAAACAGATATGA 1313
 QY 1016 TAGTGTGTTGTAATGTTCTTCTGTAAGTCTGTCTGTGTCAGAGGACAGTTGATCA 1075
 Db 1312 TAGTGTGTTGTAATGTTCTTCTGTAAGTCTGTCTGTGTCAGAGGACAGTTGATCA 1253
 QY 1076 GTGAGCTCAGAGAGAGAGTCTCTGTTCTAATGCTGTCAGAGGACACTCTCTAGT 1135
 Db 1252 GTGAGCTCAGAGAGAGAGTCTCTGTTCTAATGCTGTCAGAGGACACTCTCTAGT 1193
 QY 1136 AGGAAGAGTACACAG 1195
 Db 1192 AGGAAGAGTACACAG 1133
 QY 1196 AATGTGAAATTTAGAGCTTTCTTTCTTCTCTCATGTTCTCATGTTGTGATATAT 1255
 Db 1132 AATGTGAAATTTAGAGCTTTCTTTCTTCTCTCATGTTCTCATGTTGTGATATAT 1073
 QY 1256 TACTGATTTCAAGAGTAACTTTGTTGCTGATTAAGTAAACCGTTGTTGTTTACAT 1315
 Db 1072 TACTGATTTCAAGAGTAACTTTGTTGCTGATTAAGTAAACCGTTGTTGTTTACAT 1013
 QY 1316 CTTTGGAGAGCAGAGAAACGTTTGGAAACGATACCTTTCCAGATTCCTGATTC 1375
 Db 1012 CTTTGGAGAGCAGAGAAACGTTTGGAAACGATACCTTTCCAGATTCCTGATTC 953

Db	27876	GC	CAAA	AAAT	TAT	CC	AG	AC	CA	GG	AG	GC	CG	AG	TAC	CG	CG	AT	CC	AC	CT	CAG	CT	GA	CT	GG	27817			
Qy	743	CAC	CG	CA	CG	AG	GA	CG	CA	TT	CC	AA	AT	CA	TAC	TAC	CG	GG	AG	GA	AT	CT	TT	TAT	ACT	GT	GG	AG	GT	802
Db	27816	CAC	CG	CA	CG	AG	GA	CG	CA	TT	CC	AA	AT	CA	TAC	TAC	CG	GG	AG	GA	AT	CT	TT	TAT	ACT	GT	GG	AG	GT	27757
Qy	803	GCT	GG	TCA	CA	CA	CT	TT	CT	GG	AG	TGC	AG	CC	GA	TAC	GG	GG	TGC	GG	GA	AT	CC	CA	AG	TT	CA	862		
Db	27756	GCT	GG	T	CAG	AG	CA	TT	CT	TC	GG	AG	TGC	AG	CC	GA	TAC	GG	GG	TGC	GG	GA	AT	CC	CA	AG	TT	CA	27697	
Qy	863	TGT	TG	CT	CAG	AG	CA	TT	CA	AG	CG	GT	GC	CT	CC	CT	TG	TT	CT	AA	TG	CT	CA	AC	CA	CT	TACT	922		
Db	27696	TGT	TG	CT	CAG	AG	CA	TT	CA	AG	CG	GT	GC	CT	CC	CT	TG	TT	CT	AA	TG	CT	CA	AC	CA	CT	TACT	27637		
Qy	923	GTT	CA	TG	GC	CA	CC	GG	AA	TG	ACT	TG	GG	CC	AA	T	CAC	T	GA	TT	TG	CG	TG	AT	CC	CA	AG	GA	982	
Db	27636	GTT	CA	TG	GC	CA	CC	GG	AA	TG	ACT	TG	GG	CC	AA	T	CAC	T	GA	TT	TG	CG	TG	AT	CC	CA	AG	GA	27577	
Qy	983	CA	TT	TG	GG	ACT	CT	TT	G	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	1042		
Db	27576	CA	TT	TG	GG	ACT	CT	TT	G	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	27517		
Qy	1043	TAG	GT	T	CT	GT	GC	CA	AG	GC	AG	TT	GA	T	GA	T	GA	T	GA	T	GA	T	GA	T	GA	T	GA	1102		
Db	27516	TAG	GT	T	CT	GT	GC	CA	AG	GC	AG	TT	GA	T	GA	T	GA	T	GA	T	GA	T	GA	T	GA	T	GA	27397		
Qy	1103	CTA	AG	TG	GC	CT	GC	CA	GG	GC	CA	CT	CT	CT	ACT	TG	AG	GA	AG	AG	TAC	CA	CA	GA	GA	AG	CC	GC	1162	
Db	27436	CTA	AG	TG	GC	CT	GC	CA	GG	GC	CA	CT	CT	CT	ACT	TG	AG	GA	AG	AG	TAC	CA	CA	GA	GA	AG	CC	GC	27397	
Qy	1163	AGT	CA	GA	AG	AG	TG	TG	AA	AA	CAG	CA	CA	AT	GC	AA	TG	CG	AA	AT	TG	TG	AG	CG	TT	CC	TT	TC	1222	
Db	27396	AGT	CA	GA	AG	AG	TG	TG	AA	AA	CAG	CA	CA	AT	GC	AA	TG	CG	AA	AT	TG	TG	AG	CG	TT	CC	TT	TC	27337	
Qy	1223	TT	CC	CT	CA	TG	TT	CT	CA	TG	TT	CT	GA	TG	TAT	TAC	TG	AT	TAC	TG	AT	TAC	AG	CA	CT	TA	CT	TG	1282	
Db	27336	TT	CC	CT	CA	TG	TT	CT	CA	TG	TT	CT	GA	TG	TAT	TAC	TG	AT	TAC	TG	AT	TAC	AG	CA	CT	TA	CT	TG	27277	
Qy	1283	CG	TAT	ATA	AG	TT	TAC	CG	TT	TG	TT	TT	TA	CA	T	CT	TT	TG	GA	AG	CC	AG	GA	AG	CG	TT	GG	1342		
Db	27276	CG	TAT	ATA	AG	TT	TAC	CG	TT	TG	TT	TT	TA	CA	T	CT	TT	TG	GA	AG	CC	AG	GA	AG	CG	TT	GG	27217		
Qy	1343	AAA	AC	GT	AT	CA	CC	TT	T	CC	CA	AT	CT	CG	GA	TT	CT	CG	AC	TT	CT	TG	CA	CA	CA	CA	CT	TC	1402	

Db	26796	AGTATGGCAGTGGGAGACTAAGATGTATTAACTACCTGTGAATCATATGTTGTAGGAAA	26737
Qy	1823	AGCTGTTCCTCATGCTCAACAGCACTTGAATTCAAAGCAATGTCAAAGTGGAATAGTAGATCTG	1882
Db	26736	AGCTGTTCCCATGCTTCAACAGCACTTGAATTCAAAGCAATGTCAAAGTGGAATAGTAGATCTG	26677
Qy	1883	TGGCGATATGAGAGGGATGCAGTGCCTTCCCACTTCCTGATGCAAAATGTTATATCT	1942
Db	26676	TGGCGATATGAGAGGGATGCAGTGCCTTCCCACTTCCTGATGCAAAATGTTATATCT	2661
Qy	1943	AGGTTAAACATTTGTAAATTTTTCTTAGTGTGAATGTGTATGTCGGTAAATAGGTATTAT	2002
Db	26616	AGGTTAAACATTTGTAAATTTTTCTTAGTGTGAATGTGTATGTCGGTAAATAGGTATTAT	26557
Qy	2003	ATTTTGGCCTTTACAAATACCGTAAACAATGTTTGTTCATTTTGAATACTTAATGCCAAGTAA	2062
Db	26556	ATTTTGGCCTTTACAAATACCGTAAACAATGTTTGTTCATTTTGAATACTTAATGCCAAGTAA	26497
Qy	2063	CAATGCATGCTTTTGGAAATTTGGAAGATGGTTTTATTCTTTGAGAAGCAAAATATGTTTGC	2122
Db	26496	CAATGCATGCTTTTGGAAATTTGGAAGATGGTTTTATTCTTTGAGAAGCAAAATATGTTTGC	26437
Qy	2123	ATTTAAATGCTTTGATTTGTTGCATATCAAGAAATGATTGAACGTTCTCAAAACCCCTGTTTAC	2182
Db	26436	ATTTAAATGCTTTGATTTGTTGCATATCAAGAAATGATTGAACGTTCTCAAAACCCCTGTTTAC	26377
Qy	2183	GGTACTTGGTAAAGAGGGAGCGCGTTTGGGAGAGACCAATGCATCGCTGTCCAAAGTGTTTC	2242
Db	26376	GGTACTTGGTAAAGAGGGAGCGCGTTTGGGAGAGACCAATGCATCGCTGTCCAAAGTGTTTC	26317
Qy	2243	TTGTTAAGTGCCTTTTAACTCGGAGAGGCTAACTCCTCAAAATACTTTTTTTAACTGCATTCT	2302
Db	26316	TTGTTAAGTGCCTTTTAACTCGGAGAGGCTAACTCCTCAAAATATTTTTTTTAACTGCATTCT	26257
Qy	2303	ATTAATAAATGGGCACAGTAGTATGCTCCTTAC	2331
Db	26256	ATTAATAAATGGGCACAGTAGTATGCTCCTTAC	26228

Search completed: December 14, 2002, 17:10:26
Job time : 4511.35 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2002, 15:06:10 ; Search time 662.336 Seconds
(without alignments)
14597.874 Million cell updates/sec

Title: US-09-782-953-5

Perfect score: 597

Sequence: 1 atggatttttaggacttttag.....caccgatccaccttagctga 597

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpi:*

7: em_estro:*

8: em_estc:*

9: gb_estl:*

10: gb_est2:*

11: gb_estc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: gb_gss:*

18: em_gss_hum:*

19: em_gss_inv:*

20: em_gss_pin:*

21: em_gss_vrt:*

22: em_gss_fun:*

23: em_gss_mam:*

24: em_gss_mus:*

25: em_gss_Other:*

26: em_gss_pro:*

27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	587.4	98.4	1158	14	BQ895506
2	566.6	94.9	753	13	BI148584
3	532.6	89.2	710	10	BB617325
4	532.2	89.1	665	13	BI219142
5	532	89.1	950	12	BF782633
6	530	88.8	980	12	BQ296537

7	517	86.6	711	13	BG915568
8	516	86.4	827	13	BI102432
9	506.2	84.8	1507	11	AK010696
10	504.4	84.5	559	9	AA200984
11	472.2	79.1	837	9	AU124628
12	470.6	78.8	939	9	AL546617
13	470.6	78.8	1041	13	BM450020
14	469.6	78.7	718	9	AU131040
15	469	78.6	931	9	AL543576
16	467.4	78.3	946	14	BQ278576
17	460.6	77.2	740	13	BI463566
18	459.6	77.0	885	9	AL544755
19	456.2	76.4	520	14	BQ749142
20	455.8	76.3	906	9	AL544313
21	455.4	76.3	939	9	AL554686
22	450.2	75.4	662	13	BI464521
23	445.4	74.6	1078	13	BM541636
24	441.8	74.0	769	13	BI767955
25	434.4	72.8	501	12	BF385073
26	433.8	72.7	710	12	BG475986
27	432.6	72.5	811	12	BG570239
28	431.4	72.3	751	9	AL576189
29	408.2	68.4	890	9	AL538796
30	407.8	68.3	828	9	AL536447
31	406.2	68.0	956	9	AL556803
32	404.2	67.7	705	9	AL550372
33	403	67.5	1047	12	BG287042
34	401.6	67.3	1014	9	AL559594
35	401.4	67.2	627	13	BI859506
36	399.2	66.9	875	12	BG574693
37	397.8	66.6	937	12	BE795722
38	395.6	66.3	867	14	BQ427531
39	395	66.2	923	9	AL551657
40	382	64.0	986	12	BG828534
41	375.8	62.9	859	12	BF203629
42	375.6	62.9	535	14	BM832622
43	374	62.6	822	12	BE895008
44	372.2	62.3	525	9	AL037245
45	369.4	61.9	727	13	BI327875

ALIGNMENTS

RESULT 1
BQ895506
LOCUS BQ895506 1158 bp mRNA linear EST 16-AUG-2002
DEFINITION AGENCOURT_8762889 NIH_MGC_130 Mus musculus cDNA clone IMAGE:6329174
S., mRNA sequence.
ACCESSION BQ895506
VERSION BQ895506.1 GI:22287520
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1158)
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman, Ph.D.
cDNA Library Preparation: ResGen, Invitrogen Corp
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM13781 row: i column: 15
High quality sequence start: 83
High quality sequence stop: 577.

BG915568 602815815
BI102432 602889619
AK010696 Mus muscu
AA200984 mul2e11.r
AU124628 AU124628
AL546617 AL546617
BM450020 AGENCOURT
AU131040 AU131040
AL543576 AL543576
BQ278576 AGENCOURT
BI463566 603206210
AL544755 AL544755
BQ749142 UI-M-P00-
AL544313 AL544313
AL554686 AL554686
BI464521 603206117
BM541636 AGENCOURT
BI767955 603061320
BF385073 602045671
BG475986 602521111
BG570239 602590641
AL576189 AL576189
AL538796 AL538796
AL536447 AL536447
AL556803 AL556803
AL550372 AL550372
BG287042 602382338
AL559594 AL559594
BI859506 603385358
BG574693 602596709
BE795722 601590730
BQ427531 AGENCOURT
AL551657 AL551657
BG28534 602752653
BF203629 601865655
BM832622 K-EST0107
BE895008 601433720
AL037245 DKFP2564P
BI327875 602979461

FEATURES

Location/Qualifiers
1. .1158
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_image:6329174"
/clone_1lb="NIH_MGC_130"
/lab_host="DH10B (phage-resistant)"
/note="Organ: oocytes; Vector: pCMV-SPORT6.1.ccd; Site 1: EcoRV; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.95 kb. Constructed by Resgen, Invitrogen Corp. Note: this is a NIH_MGC Library."

BASE COUNT 289 a 354 c 288 g 226 t 1 others
ORIGIN

Query Match 98.4%; Score 587.4; DB 14; Length 1158;
Best Local Similarity 99.0%; Pred. No. 6.3e-154;
Matches 591; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

1 ATGATTTTAAAGGACTTTAGTACATATTTAGCTCCCTGATTCCTGCTGGCAACGAT 60
109 ATGCATTTTAAAGGACTTTAGTACATATTTAGCTCCCTGATTCCTGCTGGCAACGAT 168
61 GATGCTTTAGGCAAGTGTAGACAGGCGCAATTGAATCCCTCTTCAAGACATATGAC 120
169 GATGCTTTAGGCAAGTGTAGACAGGCGCAATTGAATCCCTCTTCAAGACATATGAC 228
121 AAGGACACCACTTCCAGTATTTTAAAGCTTCAACGTCGCGATTAACCTTACGCAAC 180
229 AAGGACACCACTTCCAGTATTTTAAAGCTTCAACGTCGCGATTAACCTTACGCAAC 288
181 CCTTATCTGACAGCCGATGCGAGCTCCGCTCCCAAGACGAGTTCCTGGGGAAGAA 240
289 CCTTATCTGACAGCCGATGCGAGCTCCGCTCCCAAGACGAGTTCCTGGGGAAGAA 348
241 ATGAAGTTGATTTTGTCTCAGACTTTACATAGAAAGTTACACCTGCTCCGCCAAT 300
349 ATGAAGTTGATTTTGTCTCAGACTTTACATAGAAAGTTACACCTGCTCCGCCAAT 408
301 CCCGACAAACAGTTCCTGATTCCTCCCTCCGCTCTCTCCCTGCTGGTGGAAACAAGTA 360
409 CCCGACAAACAGTTCCTGATTCCTCCCTCCGCTCTCTCCCTGCTGGTGGAAACAAGTA 468
361 GAAGATGCCACCCCGTCATTAATTAAGATTTTATATGCAATCTTCAAGCTGGGGCCA 420
469 GAAGATGCCACCCCGTCATTAATTAAGATTTTATATGCAATCTTCAAGCTGGGGCCA 528
421 GGAGAGAGTATGAACTGCAATGACGAGCAGACCCCACTCCCAATGCTGGTGGTCAAGTG 480
529 GGAGAGAGTATGAACTGCAATGACGAGCAGACCCCACTCCCAATGCTGGTGGTCAAGTG 588
481 TGTGAGAGTACCAAGAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
589 TGTGAGAGTACCAAGAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 648
541 AAGCCCAAAATCATTCAGACAGGAGACCGAGTACACACCGATCCACCTTAAGCTGA 597
649 AAGCCCAAAATCATTCAGACAGGAGACCGAGTACACACCGATCCACCTTAAGCTGA 705

RESULT 2
B1148584 753 bp mRNA linear EST 05-JUL-2001
LOCUS B1148584
DEFINITION 602911955P1 NCI_GAP_L19 Mus musculus cDNA IMAGE:5053246 5',
mRNA sequence.
ACCESSION B1148584
VERSION B1148584.1 GI:14608585
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 753)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strusberg, Ph.D.
Email: cgepds-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: L14M1114 row: m column: 23
High quality sequence stop: 751.

FEATURES
source Location/Qualifiers
1. .753
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone_image:5053246"
/clone_1lb="NCI_GAP_L19"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.9 kb. Constructed by Life Technologies. Note: this is a NCI_GAP Library."

BASE COUNT 207 a 197 c 192 g 157 t
ORIGIN

Query Match 94.9%; Score 566.6; DB 13; Length 753;
Best Local Similarity 99.0%; Pred. No. 3.5e-148;
Matches 591; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

1 ATGATTTTAAAGGACTTTAGTACATATTTAGCTCCCTGATTCCTGCTGGCAACGAT 60
89 ATGCATTTTAAAGGACTTTAGTACATATTTAGCTCCCTGATTCCTGCTGGCAACGAT 148
61 GATGCTTTAGGCAAGTGTAGACAGGCGCAATTGAATCCCTCTTCAAGACATATGAC 120
149 GATGCTTTAGGCAAGTGTAGACAGGCGCAATTGAATCCCTCTTCAAGACATATGAC 207
121 AAGGACACCACTTCCAGTATTTTAAAGCTTCAACGTCGCGATTAACCTTACGCAAC 180
208 AAGGACACCACTTCCAGTATTTTAAAGCTTCAACGTCGCGATTAACCTTACGCAAC 267
181 CCTTATCTGACAGCCGATGCGAGCTCCGCTCCCAAGACGAGTTCCTGGGGAAGAA 240
268 CCTTATCTGACAGCCGATGCGAGCTCCGCTCCCAAGACGAGTTCCTGGGGAAGAA 327
241 ATGAAGTTGATTTTGTCTCAGACTTTACATAGAAAGTTACACCTGCTCCGCCAAT 300
328 ATGAAG-TGTATTTTGTCTCAGACTTTACATAGAAAGTTACACCTGCTCCGCCAAT 386
301 CCCGACAAACAGTTCCTGATTCCTCCCTCCGCTCTCTCCCTGCTGGTGGAAACAAGTA 360
387 CCCGACAAACAGTTCCTGATTCCTCCCTCCGCTCTCTCCCTGCTGGTGGAAACAAGTA 446
361 GAAGATGCCACCCCGTCATTAATTAAGATTTTATATGCAATCTTCAAGCTGGGGCCA 420
447 GAAGATGCCACCCCGTCATTAATTAAGATTTTATATGCAATCTTCAAGCTGGGGCCA 506
421 GGAGAGAGTATGAACTGCAATGACGAGCAGACCCCACTCCCAATGCTGGTGGTCAAGTG 480
507 GGAGAGAGTATGAACTGCAATGACGAGCAGACCCCACTCCCAATGCTGGTGGTCAAGTG 566
481 TGTGAGAGTACCAAGAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
567 TGTGAGAGTACCAAGAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 626
541 AAGCCCAAAATCATTCAGACAGGAGACCGAGTACACACCGATCCACCTTAAGCTGA 597
627 AAGCCCAAAATCATTCAGACAGGAGACCGAGTACACACCGATCCACCTTAAGCTGA 683

REFERENCE 1 (bases 1 to 665)
 AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: LLM1238 row: 0 column: 21
 High quality sequence stop: 657.
 Location/Qualifiers

FEATURES
 source
 1..665
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:5099372"
 /lab_host="NCI CGAP L19"
 /note="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.9 kb. Constructed by Life Technologies. Note: this is a NCI CGAP Library."
 Location/Qualifiers

BASE COUNT 192 a 170 c 158 g 145 t
 ORIGIN

Query Match 89.1%; Score 532.2; DB 13; Length 665;
 Best Local Similarity 97.3%; Pred. No. 1.5e-138; Indels 3; Gaps 3;
 Matches 573; Conservative 0; Mismatches 13;

1 ATGATTTTGGGACTTAACTAATTTAGCTCCCTGATGCTGTTGGCAACGAT 60
 77 ATGATTTTGGGACTTAACTAATTTAGCTCCCTGATGCTGTTGGCAACGAT 136
 61 GATGCTTTCAGGAAAGTGAAGCCAGGCGCAAAATTTGAATCCCTCTTACAGCATATGAC 120
 137 GATGCTTTCAGGAAAGTGAAGCCAGGCGCAAAATTTGAATCCCTCTTACAGCATATGAC 195
 121 AAGGACCCAGCTTCCAGTATTTAAGAGCTTCAACGTCTCCGATMAAATTGACGAC 180
 196 AAGGACCCAGCTTCCAGTATTTAAGAGCTTCAACGTCTCCGATMAAATTGACGAC 255
 181 CCCTTATCTGACGCGATGCGAGCTGCGCTGCAAAAGCCGAGTTCTTGGGAAAGAA 240
 256 CCCTTATCTGACGCGATGCGAGCTGCGCTGCAAAAGCCGAGTTCTTGGGAAAGAA 315
 241 ATGAAGTTGATTTTGTCTCAGACTTACACATAGGAAGTTTCAACCTGCTCCGCCAAT 300
 316 ATGAAGTTGATTTTGTCTCAGACTTACACATAGGAAGTTTCAACCTGCTCCGCCAAT 375
 301 CCCGACAAAGTTCTCTATCTCCCTCCGCTCTCTCTCCGTTGGTGGAAACAATA 360
 376 CCCGACAAAGTTCTCTATCTCCCTCCGCTCTCTCTCCGTTGGTGGAAACAATA 435
 361 GAAGATGCCACCCCGCTCATTAATTAAGATCTTTTATATGCAATGCCAGTGGGGCCA 420
 436 GACCAATTTATACCCCGCTCATTAATTAAGATCTTTTATATGCAATGCCAGTGGGGCCA 495
 421 GGAGGAAGTATGAAGTGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
 496 GGAGGAAGTATGAAGTGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 555
 481 TGTGAGAGTGAAG 539
 556 TGTGAGAGTGAAG 615
 540 CAAGCCCAAAATATATCAGACA-CGAGAGCCGAGTACACCGATCCA 587

Db 616 AAGCCCAAAATATATCAGACACCGAGAGCCGAGTACACCAATCCA 664

RESULT 5
 LOCUS BF782633 950 bp mRNA linear EST 12-JAN-2001
 DEFINITION 602107462P1 NCI CGAP Kid14 Mus musculus cDNA clone IMAGE:4236038
 5', mRNA sequence.
 ACCESSION BF782633
 VERSION BF782633.1 GI:12087669
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 950)
 AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: LLM9843 row: k column: 15
 High quality sequence stop: 701.
 Location/Qualifiers

FEATURES
 source
 1..950
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:4236038"
 /lab_host="NCI CGAP Kid14"
 /note="Organ: kidney; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.75 kb. Constructed by Life Technologies. Note: this is a NCI CGAP Library."
 Location/Qualifiers

BASE COUNT 220 a 257 c 274 g 199 t
 ORIGIN

Query Match 89.1%; Score 532; DB 12; Length 950;
 Best Local Similarity 98.3%; Pred. No. 2.1e-138; Indels 5; Gaps 5;
 Matches 590; Conservative 0; Mismatches 5;

1 ATGATTTTGGGACTTAACTAATTTAGCTCCCTGATGCTGTTGGCAACGAT 60
 76 ATGATTTTGGGACTTAACTAATTTAGCTCCCTGATGCTGTTGGCAACGAT 134
 61 GATGCTTTCAGGAAAGTGAAGCCAGGCGCAAAATTTGAATCCCTCTTACAGCATATGAC 120
 135 GATGCTTTCAGGAAAGTGAAGCCAGGCGCAAAATTTGAATCCCTCTTACAGCATATGAC 194
 121 AAGGACCCAGCTTCCAGTATTTAAGAGCTTCAACGTCTCCGATMAAATTGACGAC 180
 195 AAGGACCCAGCTTCCAGTATTTAAGAGCTTCAACGTCTCCGATMAAATTGACGAC 254
 181 CCCTTATCTGACGCGATGCGAGCTGCGCTGCAAAAGCCGAGTTCTTGGGAAAGAA 240
 255 CCCTTATCTGACGCGATGCGAGCTGCGCTGCAAAAGCCGAGTTCTTGGGAAAGAA 314
 241 ATGAAGTTGATTTTGTCTCAGACTTACACATAGGAAGTTTCAACCTGCTCCGCCAAT 300
 315 ATGAAGTTGATTTTGTCTCAGACTTACACATAGGAAGTTTCAACCTGCTCCGCCAAT 373
 301 CCCGACAAAGTTCTCTATCTCCCTCCGCTCTCTCTCCGTTGGTGGAAACAATA 360
 374 CCCGACAAAGTTCTCTATCTCCCTCCGCTCTCTCTCCGTTGGTGGAAACAATA 433

```

QY 361 GAAGATGCCACCCCGTCATAAATTAGGATCTTTTATATGCCATCTCCAGCTGGGGCCA 420
Db 434 GAAGATGCCACCCCGTCATAAATTAGGATCTTTTATATGCCATCTCCAGCTGGGGCCA 493
QY 421 GGAGAGAAGTATGAACCTGCATGTCAGGACAGACCCCACTCCCAAGT-GTGGTGGTCCACGT 479
Db 494 GGAGAGAAGTATGAACCTGCATGTCAGGACAGACCCCACTCCCAAGTGGTGGTCCACGT 553
QY 480 GTGTGAGTGAACAAGAGAGATGAGGAGGAAGAGAA-GAGATGGAGAGAATG-AAGAGA 537
Db 554 GTGTGAGTGAACAAGAGAGATGAGGAGGAAGAGAAAGAGATGGAGAGAATGAAGAGA 613
QY 538 CCAAGGCCCAAAATCATCCAGACACGAGACCGGAGTACACACCGATCCACTTACGTGA 597
Db 614 CCAAGGCCCAAAATCATCCAGACACGAGACCGGAGTACACACCGATCCCACTCAGCTGA 673

RESULT 6
BG296537
LOCUS
DEFINITION
602394441F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:4506130 5',
mRNA sequence.
ACCESSION
BG296537
VERSION
BG296537.1 GI:13059271
KEYWORDS
EST.
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 980)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM10381 row: a column: 11
High quality sequence start: 3
High quality sequence stop: 747.
FEATURES
Location/Qualifiers
1..980
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:4506130"
/clone_lib="NIH_MGC_94"
/tissue_type="retina"
/lab_host="DH10B (phage-resistant)"
/Note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 3.3 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 265 a 258 c 269 g 188 t
ORIGIN

Query Match 88.8%; Score 530; DB 12; Length 980;
Best Local Similarity 98.3%; Pred. No. 7.7e-138;
Matches 578; Conservative 0; Mismatches 5; Indels 5; Gaps 4;

QY 1 ATGATTTTATGGGACCTTAGCTACATTTTAGCTCCCTGATTTGCTGGCAACGAT 60
Db 86 ATGATTTTATGGGACCTTAGCTACATTTTAGCTCCCTGA-TGCTGTGTGCAACGAT 144
QY 61 GATGCTTTCAGCGAAAGTGAGACCGGCGCAAAATTTGAATTCCTCTTCAGACATATGAC 120
Db 145 GATGCTTTCAGCGAAAGTGAGACCGGCGCAAAATTTGAATTCCTCTTCAGACATATGAC 204

```

```

QY 121 AAGGACACCACTTCCAGTATTTTAAAGAGCTTCAAAAGTGTCCGGATAAACTTTCAGCAAC 180
Db 205 AAGGACACCACTTCCAGTATTTTAAAGAGCTTCAAAAGTGTCCGGATAAACTTTCAGCAAC 264
QY 181 CCTTATCTCGAGCCGATGCCAGGCTGCGGCTGCACAAGACCGAGTTCCTGGGGAAGGAA 240
Db 265 CCTTATCTCGAGCCGATGCCAGGCTGCGGCTGCACAAGACCGAGTTCCTGGGGAAGGAA 324
QY 241 ATGAAGTTGTTATTTTGTCTCAGACTTTTACACATAGGAAGTTTACACCTGGCTCCGCCCAAT 300
Db 325 ATGAAGTTGTA-TTTTGTCTCAGACTTTTACACATAGGAAGTTTACACCTGGCTCCGCCCAAT 383
QY 301 CCGGACAAACAGTTCCTCATCTCCCTCCGCTCTCTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 360
Db 384 CCGGACAAACAGTTCCTCATCTCCCTCCGCTCTCTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 443
QY 361 GAAGATGCCACCCCGTCATAAATTAGGATCTTTTATATGCCATCTCCAGCTGGGGCCA 420
Db 444 GAAGATGCCACCCCGTCATAAATTAGGATCTTTTATATGCCATCTCCAGCTGGGGCCA 503
QY 421 GGAGAGAAGTATGAACCTGCATGTCAGGACAGACCCCACTCCCAAGTGTGGTCCACCGTG 480
Db 504 GGAGAGAAGTATGAACCTGCATGTCAGGACAGACCCCACTCCCAAGTGTGGTCCACCGTG 563
QY 481 TGTGAGAGTGACCAAGAGAGATGAGGAGGAGAGAGATGAGAGAGATGAAGAGACCC 540
Db 564 TGTGAGAGTGACCAAGAGAGATGAGGAGGAGAGAGATGAGAGAGATGAAGAGACCC 622
QY 541 AAGCCCAAAATCATCCAGACACGAGACCGGAGTACACACCGATCCAC 588
Db 623 -AAGCCCAAAATCATCCAGACACGAGACCGGAGTACACACCGATCCAC 668

RESULT 7
BG155568
LOCUS
DEFINITION
602815815F1 NCI_CGAP_Mam4 Mus musculus cDNA clone IMAGE:4945346 5',
mRNA sequence.
ACCESSION
BG155568
VERSION
BG155568.1 GI:14296044
KEYWORDS
EST.
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 711)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth
Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM10891 row: n column: 03
High quality sequence stop: 711.
FEATURES
Location/Qualifiers
1..711
/organism="Mus musculus"
/strain="NMRI"
/db_xref="taxon:10090"
/clone="IMAGE:4945346"
/clone_lib="NCI CGAP Mam4"
/tissue_type="tumor, gross tissue"
/dev_stage="5 months"
/lab_host="DH10B"
/Note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT."

```

Library constructed by Life Technologies. Investigators providing samples: Lohar Hemmighausen/Frischella Furch, NIH Reference for transgenic model: Li et al., Cell Growth and Differentiation 7, 3-11 (1996)."

BASE COUNT 196 a 184 c 183 g 148 t

Query Match 86.6%; Score 517; DB 13; Length 711;
Best Local Similarity 98.6%; Pred. No. 2,96-134;
Matches 553; Conservative 0; Mismatches 5; Indels 3; Gaps 3;

22 TACATTTTACCTCCCTGCTGCTGCGCAACGATGATCTTCAGCGAAAGTAG 81
1 TACAAATTTTACCTCCCTGCTGCTGCGCAACGATGATCTTCAGCGAAAGTAG 60
82 ACCAGGCGCAATTTGAAATCCCTCTTCAGAACATATGACAGACCACTTCAGAT 141
61 ACCA-GGCCAAATTTGAAATCCCTCTTCAGAACATATGACAGACCACTTCAGAT 119
142 TTTAAGAGCTTCAACGCTGCTCGGATAAATTCTGACCAACCCCTTATTCGACGCGATG 201
120 TTTAAGAGCTTCAACGCTGCTCGGATAAATTCTGACCAACCCCTTATTCGACGCGATG 179
202 AGCGTGGGCTGACAAAGCCGAGTCTCGGGAAGGAATGATTTGATTTGCTGAG 261
180 AGCGTGGGCTGACAAAGCCGAGTCTCGGGAAGGAATGATTTGATTTGCTGAG 239
262 ACTTACACATAGGAAGTTTACACCTGCTCGGCAATCCGCAACAAAGTTCTCATC 321
240 ACTTACACATAGGAAGTTTACACCTGCTCGGCAATCCGCAACAAAGTTCTCATC 299
322 TCCCTCCGCGCTCT-CCTCCGCTGGCTGGAAACAAGTAGATGCCACCCGCTCAT 380
300 TCCCTCCGCGCTCTCTCTCCGCTGGCTGGAAACAAGTAGATGCCACCCGCTCAT 359
381 AAATTCAGATCTTTTATATGCAATCTTCAGAGTGGGCGCAGAGAGATGATGATGCA 440
360 AAATTCAGATCTTTTATATGCAATCTTCAGAGTGGGCGCAGAGAGATGATGATGCA 419
441 TGCAGCAGACACCCCACTCCCACTGCTGCTGCAACCTGCTGAGAGAGACCAAGAA 500
420 TGCAGCAGACACCACTCCCACTGCTGCTGCAACCTGCTGAGAGAGACCAAGAA 479
501 TGCAGAGAGAGAGAGAGAGATGAGAGATGAGAGAGACCAAGCCCAATATCCAGAC 560
480 TGCAGAGAGAGAGAGAGAGATGAGAGATGAGAGAGACCAAG-CCAAATATCCAGAC 538
561 ACGGAGACCGGAGTACACACC 581
539 ACGGAGACCGGAGTACCAACC 559

RESULT 8 827 bp mRNA linear EST 26-JUN-2001
B1102432 602889619F1 NCI CGAP_Kid14 Mus musculus cDNA clone IMAGE:5044612
LOCUS 5', mRNA sequence.
DEFINITION

ACCESSION B1102432
VERSION B1102432.1 GI:14553325
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabre@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM1122 row: f column: 05
High quality sequence stop: 666.

FEATURES
Source 1..827
Location/Qualifiers

/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5044612"
/clone_lib="NCI CGAP Kid14"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: Kidney; Vector: pCMV-Sport6; Site 1: Not1; Site 2: Sal1; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.75 kb. Constructed by Life Technologies. Note: this is a NCI CGAP Library." Technology: 204 g 169 t

BASE COUNT 206 a 248 c 204 g 169 t

Query Match 86.4%; Score 516; DB 13; Length 827;
Best Local Similarity 97.1%; Pred. No. 6,1e-134;
Matches 579; Conservative 0; Mismatches 10; Indels 7; Gaps 5;

1 ATGATTTTAGGAGCTTACCTACATTTTACCTCCCTGATTTGCTGCGCAACGAT 60
76 ATGATTTTAGGAGCTTACCTACATTTTACCTCCCTGATTTGCTGCGCAACGAT 135
61 GATGCTTCAGCGCAAGTAGAGACAGGCGCAATTTGAATCCCTCTTCAGAACATATGAC 120
136 GATGCTTCAGCGCAAGTAGAGACCA-GGCCAAATTTGAATCCCTCTTCAGAACATATGAC 194
121 AAGACACCACTTCCAGATTTTAAAGCTTCAACGCTCCGGAATTAATCTCAGCAAC 180
195 AAGACACCACTTCCAGATTTTAAAGCTTCAACGCTCCGGAATTAATCTCAGCAAC 254
181 CCTTATTCGAGCGGAGTCCAGGCTGCGGTGCAACAGCAGATTCCTGGGAGAGAA 240
255 CCTTATTCGAGCGGAGTCCAGGCTGCGGTGCAACAGCAGATTCCTGGGAGAGAA 314
241 ATGAAGTTGATTTTGTCTCAGACTTACATAGAGAAATTCACACTGCTCGGCCAAT 300
315 ATGAAGTTGATTTTGTCTCAGACTTACATAGAGAAATTCACACTGCTCGGCCAAT 374
301 CCGGCAACAGTTCTGATCTCCCTCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360
375 CCGGCAACAGTTCTGATCTCCCTCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 434
361 GAAAGTGCACCCCGCTCATTAATTAAGATTTTATGCAATCTCCAGCTGGGCGCA 420
435 GAAAGTGCACCCCGCTCATTAATTAAGATTTTATATGCAATCTCCAGCTGGGCGCA 494
421 GGAGAGAGATGAACTGCATGACGACGACAGCCCACTCCCACTGCTGCTGCTGCTGCTG 480
495 GGAGAGAGATGAACTGCATGACGACGACAGCA-TCACAGTGTGTGTGTGTGTGTGTGT 553
481 TGTGAGAGTACCAAGAGATGAGG-AGGAAGAGAGAGATG--GAGAGAAATGAGAG 536
554 TGTGAGAGTACCAAGAGATGAGGAGAGAGAGAGAGAGATGAGAGAGATGAGAGAG 613
537 ACCCAAGCCCAAAATCATCCAGACA-CGAGAGCCGAGATACACACCATTCACCTT 591
614 ACCCAAGCCCAAAATCATCCAGACAACCGAGAGCCGAGATACACACCATTCACCTT 669

RESULT 9 1507 bp mRNA linear HTC 19-JAN-2002
AK010696
LOCUS AK010696
DEFINITION Mus musculus ES cells cDNA, RIKEN full-length enriched library, clone:2410048A02.Down syndrome critical region homolog 1 (human), full insert sequence.
ACCESSION AK010696

VERSION
KEYWORDS
SOURCE

AK010696.1 GI:12846329
HTC: CAP trapper.
Mus musculus (strain: C57BL/6J) ES cells cDNA to mRNA,
clone: 11b:RIKEN full-length enriched mouse cDNA library
clone: 2410048A02.

ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuina, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Onoda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
Kuhl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,
Quackenbush, J., Schriml, L.M., Stauber, J., Suzuki, R., Tomita, M.,
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C.,
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,
Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H.,
Toyo-oka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, D.,
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S.
and Hayashizaki, Y.
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
11217851

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

5 (bases 1 to 1507)
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,
Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C.,
Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T.,
Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Hori, P.,
Hume, D., Inotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koyama, T.,
Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K.,
Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J.,
Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,
Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I.,
Yasuniishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and

FEATURES
source
Location/Qualifiers
1..1507
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:2410048A02"
/db_xref="MGD:MGI:1898605"
/db_xref="taxon:10090"
/clone="2410048A02"
/cell_type="ES cells"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
1..1507
/gene="Dscr1"
208..564
/gene="Dscr1"
/note="Down syndrome critical region homolog 1 (human)"
data source:MGD, source key:MGI:1890564, evidence:ISS
putative"
/codon_start=1
/protein_id="BAB27128.1"
/db_xref="GI:12846330"
/db_xref="MGD:MGI:1890564"
/translation="MKLYFAQLHIGSHLAPPNDKQLISPPASPPVGVKQVEDAT
PVINYDULFAISKLGPGEKYEHLHAATDTTPSVVVHVHVCSDQENEEBEMENKRPKP
KIIQTRRPYTPHLS"
BASE COUNT 354 a 402 c 390 g 361 t
ORIGIN

Query Match 84.8%; Score 506.2; DB 11; Length 1507;
Best Local Similarity 99.4%; Pred. No. 4.8e-131;
Matches 508; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 87 GGCCAAATTGGAATCCCTCTTCAGAACATATGACAGGACACACCTTCAGTATTATTA 146
Db 54 GGCCAAATTGGAATCCCTCTTCAGAACATATGACAGGACACACCTTCAGTATTATTA 113

QY 147 GAGCTTCAAACGTCGCGATAAATCTTCAGAACACCTTCAGTATTATTCGAGGATG 206
Db 114 GAGCTTCAAACGTCGCGATAAATCTTCAGAACACCTTCAGTATTATTCGAGGATG 173

QY 207 CGCGCTGCACAAACGAGTTCCTCGGGAGGAAGAAATGAAGTTGTATTTCGTCAGACTTT 266
Db 174 CGCGCTGCACAAACGAGTTCCTCGGGAGGAAGAAATGAAGTTGTATTTCGTCAGACTTT 233

QY 267 ACACATAGGAAGTTACACCTCGCTCCGCCCAATCCGACAAACAGTTCCTCATCTCCCC 326
Db 234 ACACATAGGAAGTTACACCTCGCTCCGCCCAATCCGACAAACAGTTCCTCATCTCCCC 293

QY 327 TC CGCGCTCTCTCCCGTGGCTGGAGAAACAGTAGAAGATGCCACCCCGGTCAATAATTA 386

COMMENT
Direct Submission
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
Please visit our web site (http://genome.gsc.riken.go.jp/) for
further details.
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in Riken.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues. First strand cDNA was primed with a primer
[5' GAGAGAGAGATCCAGAGCTCTTTTCTTTTNN 3'], cDNA was
prepared by using triose thermo-activated reverse transcriptase
and subsequently enriched for full-length by cap-trapper. cDNA went
through one round of normalization to Rot = 5.0 and subtraction to
Rot = 25.0. Second strand cDNA was prepared with the primer adapter
of sequence 15'
GAGAGAGATTCGAGTTAATTAATATCCCTCCCCCCCC 3'. cDNA was cleaved
with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.
Host: SOLR.

```

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:639212"
/clone_lib="Soares_thymus_2NbMT"
/sex="male"
/tissue_type="Thymus"
/dev_stage="4 weeks"
/lab_host="DH10B"
/notes="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5',
TGTTCACATCTGGAAGTGGAGCGCGCGCTTTTCTTTTCTTTTCTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not
I and Eco RI sites of the modified pT7T3 vector. RNA
provided by Dr. Bertrand Jordan. Library went through two
rounds of normalization, and was constructed by Bento

```

Tel.: 81-438-52-3986
 Fax: 81-438-52-3986
 Email: genomic@hri.co.jp
 HRI human cDNA project, 5'- & 3'-end one pass sequencing: Helix

```

*
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
Location/Qualifiers
1. .837
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NT2RM4000307"
/cell_type="teratocarcinoma"
/clone_lib="NT2RM4"
/cell_line="NT2"
/note="Vector: pME18SPL3; mRNA from uninduced NT2 neuronal
precursor cells"
BASE COUNT 241 a 201 c 207 g 186 t 2 others
ORIGIN
Query Match 79.1%; Score 472.2; DB 9; Length 837;
Best Local Similarity 88.1%; Pred. No. 1.2e-121;
Matches 526; Conservative 0; Mismatches 68; Indels 3; Gaps 1;
QY 1 ATGGATTTTGGGACCTTAGCTACAAATTTAGCTCCCTGATTTGTTGTGGCAAAACGAT 60
Db 118 ATGCATTTTGAACATTTTAACTACAGTTTATGCTCCCTGATGCTGTGGCAACAGT 177
QY 61 GATGCTTTCAGCGAAAGTCAGACACGAGGCGCAAAATTTGAATCCCTCTCAGAACATATGAC 120
Db 178 GATATCTTCAGCGAAAGTGAACACGAGGCGCAAAATTTGATGCTCCCTCTTAGGACGTATGAC 237
QY 121 AAGGACACCACTTCAGTATTTTAAAGCTTTCAAACTGTCCGGATAACTTCAGCAAC 180
Db 238 AAGGACATCACCTTTTCAGTATTTTAAAGCTTTCAAACTGTCCGGATAACTTCAGCAAC 297
QY 181 CCCTTATCTGCAGCGGATGCCAGGCTGCGGTGCACAGACGAGTTCCTGGGAAAGGAA 240
Db 298 CCCTTCTCGCAGCAGATGCCAGGCTCAGCTGCATAGACTGATTTCTGGGAAAGGAA 357
QY 241 ATGAAGTTGTTATTTTGTCTCAGACTTTTACATAGCAAGTTTCAACCTGGCTCCGCCCAAT 300
Db 358 ATGAAGTTTATATTTTGTCTCAGACTTTTACATAGCAAGTTTCAACCTGGCTCCGCCCAAT 417
QY 301 CCGGCAACAAGTTTCTCATCTCCCTCCGCTCCGCTCTCTCCGTTGGCTGGAAACAAGTA 360
Db 418 CCAGCAACAGCTTTCTGATCTCCCTCCGCTCTCCGCTCCGCTGGGATGGAAACAAGTG 477
QY 361 GAAGATGCCACCCCGCTCATATAATTTAGCTTTTATATGCCATCTCCAAGCTTGGGCCA 420
Db 478 GAAGATGCCACCCCGCTCATATAATTTAGCTTTTATATGCCATCTCCAAGCTTGGGCCA 537
QY 421 GGAGAGATGATGAATGCAATGCAGACAGACCCCACTCCAGTGTGGTGGTCCACGTG 480
Db 538 GGGGAAAGTATGAATTTGCATGCAGCGACTGCACCACTCCAGCTGGTGGTCCATGTA 597
QY 481 TGTGAGATGACCAAGAGAATGAGGAGAGAGAGAGATGAGAGAGATGAAGAGACCC 540
Db 598 TGTGAGATGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 654
QY 541 AAGCCCAAAATCATCCAGACACCGAGACCGGAGTACACACCGATCCACCTTAGCTGA 597
Db 655 AAGCCCAAAATTTATCCAGACACCGAGGCGGAGTACACCGGATCCACCTTAGCTGA 711

RESULT 12
AL546617 939 bp mRNA linear EST 16-FEB-2001
LOCUS AL546617 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0D1029J21 5
DEFINITION prime, mRNA sequence.
ACCESSION AL546617
VERSION AL546617.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

REFERENCE AUTHORS TITLE JOURNAL COMMENT

1 (bases 1 to 939)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1. .939
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1029J21"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

FEATURES source

BASE COUNT 263 a 219 c 238 g 218 t 1 others
ORIGIN
Query Match 78.8%; Score 470.6; DB 9; Length 939;
Best Local Similarity 87.9%; Pred. No. 3.7e-121;
Matches 525; Conservative 0; Mismatches 69; Indels 3; Gaps 1;
QY 1 ATGGATTTTGGGACCTTTAGCTACAAATTTAGCTCCCTGATTTGTTGTGGCAAAACGAT 60
Db 53 ATGCATTTTGAACATTTTAACTACAGTTTATGCTCCCTGATTTGCTGTGGCAACAGT 112
QY 61 GATGCTTTCAGCGAAAGTCAGACACGAGGCGCAAAATTTGAATCCCTCTCAGAACATATGAC 120
Db 113 GATATCTTCAGCGAAAGTGAACACGAGGCGCAAAATTTGATGCTCCCTCTTAGGACGTATGAC 172
QY 121 AAGGACACCACTTCAGTATTTTAAAGCTTTCAAACTGTCCGGATAACTTCAGCAAC 180
Db 173 AAGGACATCACCTTTTCAGTATTTTAAAGCTTTCAAACTGTCCGGATAACTTCAGCAAC 232
QY 181 CCCTTATCTGCAGCGATGCCAGGCTGCGGTGCACAGACCGAGTTCCTGGGAAAGGAA 240
Db 233 CCCTTCTCGCAGCAGATGCCAGGCTCAGCTGCATAGACTGATTTCTGGGAAAGGAA 292
QY 241 ATGAAGTTGTTATTTTGTCTCAGACTTTTACATAGCAAGTTTCAACCTGGCTGGGAAAGGAA 300
Db 293 ATGAAGTTTATATTTTGTCTCAGACTTTTACATAGCAAGTTTCAACCTGGCTGGGAAAGGAA 352
QY 301 CCGGCAACAAGTTTCTCATCTCCCTCCGCTCTCTCCCTGGTGGTGGTCCACGTG 360
Db 353 CCAGCAAGCAGATTTCTGATCTCCCTCCGCTCTCTCCGCTGGGATGGGAAAGGAA 412
QY 361 GAAGATGCCACCCCGCTCATATAATTTAGCTTTTATATGCCATCTCCAAGCTTGGGCCA 420
Db 413 GAAGATGGACCCCACTCATATAATTTAGCTTTTATATGCCATCTCCAAGCTTGGGCCA 472
QY 421 GGAGAGATGATGAATGCAATGCAGGACAGACCCCACTCCAGTGTGGTGGTCCACGTG 480
Db 473 GGGGAAAGTATGAATTTGCAACGCGAGCTGCACCACTCCAGCTGGTGGTCCATGTA 532
QY 481 TGTGAGATGACCAAGAGAATGAGGAGAGAGAGAGAGATGAGAGAGATGAAGAGACCC 540
Db 533 TGTGAGATGATCAAG 589
QY 541 AAGCCCAAAATCATCCAGACACCGAGACCGGAGTACACACCGATCCACCTTAGCTGA 597
Db 590 AAGCCCAAAATTTATCCAGACACCGAGGCGGAGTACACCGGATCCACCTTAGCTGA 646

```

RESULT 13
BM450020      1041 bp      mRNA      linear      EST 05-FEB-2002
LOCUS         AGENCOURT_6393352 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5528025
DEFINITION    5', mRNA sequence.
ACCESSION     BM450020
VERSION       BM450020.1 GI:18499060
KEYWORDS      EST.
SOURCE        human.
ORGANISM      Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE     1 (bases 1 to 1041)
AUTHORS       NIH-MGC http://mgi.nci.nih.gov/
TITLE         National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL       Unpublished (1999)
COMMENT       Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-remail.nih.gov
              Tissue Procurement: ATCC/DCTD/DTF
              DNA Library Preparation: Life Technologies, Inc.
              DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LNL at:
              http://image.jnl.gov
              Plate: LLM12204 row: d column: 10
              High quality sequence stop: 643.
              Location/Qualifiers
                1..1041
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"
                  /clone="IMAGE:5528025"
                  /clone_1db="NIH_MGC_72"
                  /tissue_type="melanotic melanoma"
                  /lab_host="DH10B (phage-resistant)"
                  /note="Organ: skin; Vector: pCMV-Sport6; Site: 1: NCI;
                  Site 2: Salt; Cloned unidirectionally. Primer: Oligo dT.
                  Average insert size 2 kb. Library constructed by Life
                  Technologies."
BASE COUNT    282 a      248 c      274 g      236 t      1 others
ORIGIN
Query Match      78.8%; Score 470.6; DB 13; Length 1041;
Best Local Similarity 87.9%; Pred. No. 3.9e-121;
Matches 525; Conservative 0; Mismatches 69; Indels 3; Gaps 1;

OY 1 ATGATTTTGAAGGACTTATTTAGTCAATTTTATGCTCCCTGATTTGTTGGGCAACGAT 60
DB 84 ATGATTTTGAAGGACTTATTTAGTCAATTTTATGCTCCCTGATTTGTTGGGCAACGAT 143
OY 61 GATGCTTCAGCGAAGGAGGACGAGGCAAAATTTGATCCCTCTTCAGAACATATGAC 120
DB 144 GATATCTTCAGCGAAGGAGGACGAGGCAAAATTTGATCCCTCTTCAGAACATATGAC 203
OY 121 AAGGACACCACTTCAGTATTTTAAAGCTTCAACGTTCCGATTAATTTAGCAAC 180
DB 204 AAGGACATCACTTCAGTATTTTAAAGCTTCAACGAGTCAGATTAATTTAGCAAC 263
OY 181 CCTTATCTGACGCGAGTCCGAGTGGGCTGCACAGACCGAGTTCTTGAGGAGGAA 240
DB 264 CCTTCTCGGAGAGATGCGAGGCTCCAGCTGCAATAGACTGAGTTCTTGAGGAGGAA 323
OY 241 ATGAGTTGATTTTGTCTGAGCTTATACATAGAGATTTCAACATGCTGCGCCCAAT 300
DB 324 ATGAGTTGATTTTGTCTGAGCTTATACATAGAGATTTCAACATGCTGCGCCCAAT 383
OY 301 CCCGACAAACAGTTCTCATCTCTCCCTCGGCTCTCTCCCTGCTGGGCTGAGAAAGTA 360
DB 384 CCAGACAAACAGTTCTCATCTCTCCCTCGGCTCTCTCCCTGCTGGGCTGAGAAAGTA 420
OY 361 GAAGATGCAACCCCTGCTTAAATTTAGATCTTTATATGATCCATCTCCAGCTGGGGCA 400
DB 444 GAAGATGCAACCCCTGCTTAAATTTAGATCTTTATATGATCCATCTCCAGCTGGGGCA 503

```

```

OY 421 GGAGAGAGATATGAACTGCATGACGACGACAGACCCCACTCCCACTGCTGCTCCAGTG 480
DB 504 GGGGAAAGATATGAACTGCATGACGACGACGACGACCACTCCCACTGCTGCTCCAGTG 563
OY 481 TGTGAGATGACCAAGAGATATGAGAGAGAGAGAGATGAGAGAGAGAGAGAGAGAGAG 540
DB 564 TGTGAGATGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 620
OY 541 AAGCCCAAAATCATCCAGACGACGAGACCGAGATACACAGCATCCTTACTGA 597
DB 621 AAGCCCAAAATCATCCAGACGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 677

RESULT 14
AU131040      718 bp      mRNA      linear      EST 01-AUG-2002
LOCUS         AU131040
DEFINITION    NT2RP3 Homo sapiens cDNA clone NT2RP3001895 5', mRNA
sequence.
ACCESSION     AU131040
VERSION       AU131040.1 GI:10991394
KEYWORDS      EST.
SOURCE        human.
ORGANISM      Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE     1 (bases 1 to 718)
AUTHORS       Oca.T., Nishikawa,T., Suzuki,Y., Ishii,S., Saio,K., Kawai,Y.,
              Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
              Isogai,T.
              HRI human cDNA project
              Unpublished (2000)
              Contact: Takao Isogai
              Genomics Laboratory
              Helix Research Institute
              1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
              Tel: 81-438-52-3975
              Fax: 81-438-52-3986
              Email: genomeshri.co.jp
              HRI human cDNA project; 5'- & 3'-end one pass sequencing; Helix
              Research Institute; cDNA library construction; Department of
              Virology, Institute of Medical Science, University of Tokyo, and
              Helix Research Institute.
              Location/Qualifiers
                1..718
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"
                  /clone="NT2RP3001895"
                  /clone_1db="NT2RP3"
                  /cell_type="teratocarcinoma"
                  /note="Vector: pME18SFL3; mRNA from NT2 neuronal precursor
                  cells after 2-weeks retinoic acid (RA) induction"
BASE COUNT    217 a      174 c      169 g      155 t      3 others
ORIGIN
Query Match      78.7%; Score 469.6; DB 9; Length 718;
Best Local Similarity 87.8%; Pred. No. 6.1e-121;
Matches 524; Conservative 0; Mismatches 70; Indels 3; Gaps 1;

OY 1 ATGATTTTGAAGGACTTATTTAGTCAATTTTATGCTCCCTGATTTGTTGGGCAACGAT 60
DB 102 ATGATTTTGAAGGACTTATTTAGTCAATTTTATGCTCCCTGATTTGTTGGGCAACGAT 161
OY 61 GATGCTTCAGCGAAGGAGGACGAGGCAAAATTTGATCCCTCTTCAGAACATATGAC 120
DB 162 GATATCTTCAGCGAAGGAGGACGAGGCAAAATTTGATCCCTCTTCAGAACATATGAC 221
OY 121 AAGGACACCACTTCAGTATTTTAAAGCTTCAACGTTCCGATTAATTTAGCAAC 180
DB 222 AAGGACATCACTTCAGTATTTTAAAGCTTCAACGAGTCAGATTAATTTAGCAAC 281
OY 181 CCTTATCTGACGCGAGTCCGAGTGGGCTGCACAAAGCGAGTTCTTGAGGAGGAA 240

```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 11:37:41 ; Search time 11.3524 Seconds
(without alignments)
513.170 Million cell updates/sec

Title: US-09-782-953-3

Perfect score: 1048

Sequence: 1 MEEVDLQDLPSTACHLDP.....RPKPKIIQTRPEVPIHLS 198

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA:**
1: /cgn2_6/ptodata/1/iaa/5A COMB.pep:**
2: /cgn2_6/ptodata/1/iaa/5B COMB.pep:**
3: /cgn2_6/ptodata/1/iaa/6A COMB.pep:**
4: /cgn2_6/ptodata/1/iaa/6B COMB.pep:**
5: /cgn2_6/ptodata/1/iaa/6C COMB.pep:**
6: /cgn2_6/ptodata/1/iaa/6D COMB.pep:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	835.5	79.7	171	2	US-08-665-040-2
2	88	8.4	284	5	PCT-US94-09752-4
3	84.5	8.1	719	4	US-09-641-741-28
4	82	7.8	1128	4	US-09-641-741-30
5	82	7.8	1128	4	US-09-060-482-8
6	81	7.7	1128	1	US-08-111-939-2
7	79	7.5	1132	1	US-08-446-038B-18
8	79	7.5	1132	1	US-08-446-010B-18
9	79	7.5	1132	2	US-08-805-445-18
10	79	7.5	1132	2	US-08-064-067D-18
11	79	7.5	1132	2	US-09-066-208-18
12	79	7.5	1142	1	US-08-057-997A-11
13	79	7.5	1142	4	US-08-665-574C-11
	79	7.5	1142	1	US-08-946-994-11
	79	7.5	1154	1	US-08-357-598-7
	79	7.5	1154	1	US-08-446-010B-24
	79	7.5	1154	2	US-09-003-289-7
	79	7.5	1154	5	PCT-US95-16435-7
	79	7.5	547	4	US-09-877-730-26
	79	7.5	624	4	US-09-877-730-24
	79	7.5	628	4	US-09-877-730-30
	79	7.5	712	4	US-09-877-730-22
	79	7.5	712	4	US-09-877-730-28
	79	7.5	712	4	US-09-877-730-16
	79	7.5	712	4	US-09-877-730-6
	79	7.5	712	4	US-09-877-730-20
	79	7.5	712	4	US-09-877-730-10

28	78.5	7.5	991	4	US-09-877-730-12
29	78.5	7.5	1069	4	US-09-877-730-2
30	78.5	7.5	1072	4	US-09-877-730-18
31	78.5	7.5	1150	4	US-09-877-730-8
32	78.5	7.5	1706	2	US-08-459-568-2
33	78.5	7.5	1706	2	US-08-399-411-2
34	78.5	7.5	1706	3	US-08-516-859A-2
35	78.5	7.5	1706	4	US-09-586-472-2
36	78.5	7.5	1706	4	US-09-528-706-2
37	77	7.3	362	2	US-08-985-090-5
38	77	7.3	362	3	US-09-165-543-32
39	77	7.3	522	6	RE34606-6
40	77	7.3	1912	4	US-08-913-832A-2
41	77	7.3	1912	4	US-09-249-181A-2
42	76	7.3	445	3	US-09-165-543-5
43	75	7.2	1158	4	US-09-060-482-2
44	73.5	7.0	560	2	US-08-928-692-53
45	73.5	7.0	560	4	US-09-339-972-53

ALIGNMENTS

RESULT 1
US-08-665-040-2
; Sequence 2, Application US/08665040
; Patent No. 5869318
; GENERAL INFORMATION:
; APPLICANT: ESTIVILL PALLEJA, XAVIER
; APPLICANT: FUTES, JUAN JOSE
; APPLICANT: PRITCHARD, MELANIE
; TITLE OF INVENTION: A NEW GENE SEQUENCE OF THE
; TITLE OF INVENTION: DOWN SYNDROM CRITICAL REGION OF HUMAN CHROMOSOME 21,
; TITLE OF INVENTION: IDENTIFIED BY A NEW "ALU-SPLICING PCR" TECHNIQUE,
; TITLE OF INVENTION: CODING FOR A PROLINE-RICH PROTEIN (DSCR1) HIGHLY
; TITLE OF INVENTION: EXPRESSED IN FETAL BRAIN AND IN HEART AND METHOD
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LADAS & PARRY
; STREET: 26 WEST 61ST STREET
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10023
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORD PERFECT 5.1 FOR DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/665,040
; FILING DATE: JUNE 7, 1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: ES P9501140
; FILING DATE: JUNE 7, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: JANET I. CORD
; REGISTRATION NUMBER: 33,778
; REFERENCE/DOCKET NUMBER: U010815-9
; TELEPHONE: (212) 708-1800
; TELEFAX: (212) 246-8959
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 171 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-665-040-2

Query Match 79.7%; Score 835.5; DB 2; Length 171;
 Best Local Similarity 94.7%; Pred. No. 8.4e-89;
 Matches 160; Conservative 5; Mismatches 3; Indels 1; Gaps 1;

30 AKFESLFRTYDKDTTQYFKSKRVRINSPNLSAADARLRLHKTFFLGKEMKLYFAOTL 89
 4 AKFESLFRTYDKDTTQYFKSKRVRINSPNLSAADARLRLHKTFFLGKEMKLYFAOTL 63
 90 HIGSSHLAPNPCKOLISPPASPPVGMQVEDATPVINVDLVAISKLGGCEKYLHAA 149
 64 HIGSSHLAPNPCKOLISPPASPPVGMQVEDATPVINVDLVAISKLGGCEKYLHAA 123

150 TDPTPVVVHVCSDOBNBEEEMERMRPKIIOTRPEYTPIHLS 198
 124 TDTTPVVVHVCSDOBNBEEEMERMRPKIIOTRPEYTPIHLS 171

RESULT 2
 PCT-US94-09752-4
 Sequence 4, Application PC/TUS9409752
 GENERAL INFORMATION:
 APPLICANT: David S. Strayer and Avinash Chander
 TITLE OF INVENTION: Compositions and Methods for
 TITLE OF INVENTION: Targeting Cells and Modulating Pulmonary Surfactant Secretion
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Jane Massey Licata, Esq.
 STREET: 210 Lake Drive East, Suite 201
 CITY: Cherry Hill
 STATE: NJ
 COUNTRY: USA
 ZIP: 08002
 COMPUTER READABLE FORM:
 MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB
 MEDIUM TYPE: STORAGE
 COMPUTER: IBM PS/2
 OPERATING SYSTEM: PC-DOS
 SOFTWARE: WORDPERFECT 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US94/09752
 FILING DATE: Herewith
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/176,218
 FILING DATE: December 30, 1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/114,951
 FILING DATE: August 31, 1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Jane Massey Licata
 REGISTRATION NUMBER: 32,257
 REFERENCE/DOCKET NUMBER: JEFF-0042
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (609) 779-2400
 TELEFAX: (609) 779-8488
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 284
 TYPE: amino acid
 TOPOLOGY: linear
 PCT-US94-09752-4

Query Match 8.4%; Score 88; DB 5; Length 284;
 Best Local Similarity 24.1%; Pred. No. 0.053;
 Matches 40; Conservative 24; Mismatches 62; Indels 40; Gaps 7;

27 LCRAKESLFRTYDKDTTQYFKSKRVRINSPNLSAADARLRLHKTFFLGKEMKLYFA 86
 3 MCGYMGSLFRSGSROTLFA-----SQMRYAD---LYASFIN---LLYYP 43
 87 QTLHIGSSHLAPNPCKOLISPPASPPVGMQVEDATPVINVDLVAISKLGGCEKYLH 135
 44 FSYLFLRAAHVLMHESVTEHTHVDINEMESPLATNRKTSVDFKD---TDYKRHQLTNSI 99

Query Match 8.1%; Score 84.5; DB 4; Length 719;
 Best Local Similarity 21.8%; Pred. No. 0.53;
 Matches 31; Conservative 27; Mismatches 37; Indels 47; Gaps 7;

136 SKLGGKELHAAATDPTPSVVHVCSDOBNBEEEMERMRPK 181
 100 SEIKPNIPLF-----APQITHCHEDEDDEEEEBEEBCGKPK 139

RESULT 3
 US-09-641-741-28
 Sequence 28, Application US/09641741
 Patent No. 6420155
 GENERAL INFORMATION:
 APPLICANT: Kerry E. Quinn
 APPLICANT: Curagen Corporation
 TITLE OF INVENTION: Aortic Carboxypeptidase-Like Proteins and Nucleic Acids
 TITLE OF INVENTION: encoding Same
 FILE REFERENCE: 15966-581
 CURRENT APPLICATION NUMBER: US/09/641,741
 PRIOR FILING DATE: 2000-08-18
 PRIOR APPLICATION NUMBER: 60/159,613
 PRIOR FILING DATE: 1999-10-14
 PRIOR APPLICATION NUMBER: 60/175,534
 PRIOR FILING DATE: 2000-01-12
 PRIOR APPLICATION NUMBER: 60/224,086
 PRIOR FILING DATE: 2000-08-09
 NUMBER OF SEQ ID NOS: 32
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO: 28
 LENGTH: 719
 TYPE: PRT
 ORGANISM: Mus musculus
 US-09-641-741-28

Query Match 8.1%; Score 84.5; DB 4; Length 719;
 Best Local Similarity 21.8%; Pred. No. 0.53;
 Matches 31; Conservative 27; Mismatches 37; Indels 47; Gaps 7;

54 VRINSPNLSAADARLRLHKTFFLGKEMKLYFAOTLHIGSSHLAPNPCKOLISPP 110
 585 LVVDSRPMTPQGRMQRRLQYRLRMEQWL---RLNSTGAPATSPV---ALMP 637
 111 ASP-----PVGMK--QVEDATPVINVDLVAISKLGGCEKYLHAAATDPT 153
 638 PSPTPAITLRPEVLPTTAGHESSETTYTEVY-----EFTEYGTD-- 681

154 PSVVHVCSDOBNBEEEMERMRPKIIOTRPEYTPIHLS 175
 682 -----LEVEIEEBEEEBEEMD 699

RESULT 4
 US-09-641-741-30
 Sequence 30, Application US/09641741
 Patent No. 6420155
 GENERAL INFORMATION:
 APPLICANT: Kerry E. Quinn
 APPLICANT: Curagen Corporation
 TITLE OF INVENTION: Aortic Carboxypeptidase-Like Proteins and Nucleic Acids
 TITLE OF INVENTION: encoding Same
 FILE REFERENCE: 15966-581
 CURRENT APPLICATION NUMBER: US/09/641,741
 PRIOR FILING DATE: 2000-08-18
 PRIOR APPLICATION NUMBER: 60/159,613
 PRIOR FILING DATE: 1999-10-14
 PRIOR APPLICATION NUMBER: 60/175,534
 PRIOR FILING DATE: 2000-01-12
 PRIOR APPLICATION NUMBER: 60/224,086
 PRIOR FILING DATE: 2000-08-09
 NUMBER OF SEQ ID NOS: 32
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO: 30
 LENGTH: 1128
 TYPE: PRT
 ORGANISM: Mus musculus

US-09-641-341-30

```

Query Match
Best Local Similarity 7.8%; Score 82; DB 4; Length 1128;
Matches 33; Conservative 29; Mismatches 37; Indels 60; Gaps 8;

QY 50 SFKVR-----INFNPLSAADARLRLHKTFF---LGKEMKLYFAQTLHIGS 93
DB 977 NWKRIREILAMNGNRPILGVDPSPMTPOQRMQORRLQYRLRMREQML---RRLNSTA 1033
QY 94 SHLAPPNDKQFLISPPASP-----PVGWK--OVEDATPVINYDLYAIS 136
DB 1034 GPATSPTP-----ALMPPSPPTAILRPWEVLPTTTAGWEESETETYTEVVT----- 1081
QY 137 KLGPEKVELHAADTPSPVVHVHVCSDQNEEEEME 175
DB 1082 -----EFETGYGTD-----LEVEIEEEEEEEEMD 1108

```

RESULT 5

```

US-09-060-482-8
; Sequence 8, Application US/09060482
; Patent No. 6458766
; GENERAL INFORMATION:
; APPLICANT: Lee, Mu-En
; APPLICANT: Layne, Matthew D.
; APPLICANT: Yet, Shaw-Fang
; TITLE OF INVENTION: AORTIC CARDOXYPEPTIDASE-LIKE POLYPEPTIDE
; FILE REFERENCE: 05433/036001
; CURRENT APPLICATION NUMBER: US/09/060,482
; EARLIER FILING DATE: 1998-04-15
; EARLIER APPLICATION NUMBER: US 08/818,009
; EARLIER FILING DATE: 1997-03-14
; EARLIER APPLICATION NUMBER: US 60/013,439
; EARLIER FILING DATE: 1996-03-15
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 1128
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-060-482-8

```

```

Query Match
Best Local Similarity 7.8%; Score 82; DB 4; Length 1128;
Matches 33; Conservative 29; Mismatches 37; Indels 60; Gaps 8;

QY 50 SFKVR-----INFNPLSAADARLRLHKTFF---LGKEMKLYFAQTLHIGS 93
DB 977 NWKRIREILAMNGNRPILGVDPSPMTPOQRMQORRLQYRLRMREQML---RRLNSTA 1033
QY 94 SHLAPPNDKQFLISPPASP-----PVGWK--OVEDATPVINYDLYAIS 136
DB 1034 GPATSPTP-----ALMPPSPPTAILRPWEVLPTTTAGWEESETETYTEVVT----- 1081
QY 137 KLGPEKVELHAADTPSPVVHVHVCSDQNEEEEME 175
DB 1082 -----EFETGYGTD-----LEVEIEEEEEEEEMD 1108

```

RESULT 6

```

US-08-111-939-2
; Sequence 2, Application US/08111939
; Patent No. 5460951
; GENERAL INFORMATION:
; APPLICANT: Kawai, Shinji
; APPLICANT: Takeshita, Sunao
; APPLICANT: Okazaki, Makoto
; APPLICANT: Amann, Egon
; TITLE OF INVENTION: Bone-Related Carboxypeptidase-Like
; OF INVENTION: Protein and Process for its Production
; SEQUENCES: 27
; ADDRESS:

```

```

; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/111,939
; FILING DATE: 26-AUG-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 324033/92
; FILING DATE: 03-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 230029/92
; FILING DATE: 28-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Forman, David S.
; REGISTRATION NUMBER: 33,694
; REFERENCE/DOCKET NUMBER: 02481.1321-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4000
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1128 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-111-939-2

Query Match
Best Local Similarity 7.7%; Score 81; DB 1; Length 1128;
Matches 33; Conservative 29; Mismatches 37; Indels 60; Gaps 8;

QY 50 SFKVR-----INFNPLSAADARLRLHKTFF---LGKEMKLYFAQTLHIGS 93
DB 977 NWKRIREILAMNGNRPILGVDPSPMTPOQRMQORRLQYRLRMREQML---RRLNSTA 1033
QY 94 SHLAPPNDKQFLISPPASP-----PVGWK--OVEDATPVINYDLYAIS 136
DB 1034 GPATSPTP-----ALMPPSPPTAILRPWEVLPTTTAGWEESETETYTEVVT----- 1081
QY 137 KLGPEKVELHAADTPSPVVHVHVCSDQNEEEEME 175
DB 1082 -----EFETGYGTD-----LEVEIEEEEEEEEMD 1108

```

RESULT 7

```

US-08-446-038B-18
; Sequence 18, Application US/08446038B
; Patent No. 5658791
; GENERAL INFORMATION:
; APPLICANT: Wilks, Andrew F.; Ziemiecki, Andrew;
; APPLICANT: Harpur, Ailsa
; TITLE OF INVENTION: No. 5658791e1 Protein Tyrosine Kinase
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM PS/2

```

OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446.038B
FILING DATE: 19-MAY-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/064.067
FILING DATE: 30-Jun-1993
APPLICATION NUMBER: PCT/US91/08889
FILING DATE: 26-NOV-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: Australian PK3594/90
FILING DATE: 28-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: Australian 88229/91
FILING DATE: 27-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5658791man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5244
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-688-9200
TELEFAX: 212-838-3884
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 1132 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-446-038B-18

Query Match 7.5%; Score 79; DB 1; Length 1132;

Best Local Similarity 24.0%; Pred. No. 4.5;
Matches 48; Conservative 30; Mismatches 64; Indels 58; Gaps 13;

12 ATIACHLDPFVFDGLCRAKFESLFTYDKDTTFQYFKSFKVRINFSNPLSAADARLRL 71
64 AAQACRISP-----LC-----HNLFLYDENTKLMWAPN--RTTVDDKMS-----LRL 105
QY 72 HKTFLQKEMKLYPAQTGHTLHSSHLAPNPDQKFLSPASPVGW--KQVEDATPVNY 129
Db 106 H-----YRMRFT-----NMHGTDNDSQSVRHSPPKQKN-GYEKKKIPDAPLDA 152
QY 130 DLVAISKLGPEKYEHLAATDPTSPVVHVCSQDENESEEE-----ME 175
Db 153 SLELYFAQG---QYDLVYKCLAP-----IRDPKTEODGHDIEHCLGMVLAISHYAMMK 204
QY 176 RMKRPK-PKIIOTRREPTP 194
Db 205 KMQLPELPKDISYKR--YIP 222

RESULT 8

US-08-446-010B-18

Sequence 18, Application US/0846010B
Patent No. 5716818
GENERAL INFORMATION:
APPLICANT: Wilks, Andrew F.; Ziemiecki, Andrew;
APPLICANT: Harpur, Alisa
TITLE OF INVENTION: No. 5716818el Protein Tyrosine Kinase
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pelte & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446.010B
FILING DATE: 19-MAY-1995
CLASSIFICATION: 433
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/446.038
FILING DATE: 19-MAY-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/064.067
FILING DATE: 30-Jun-1993
APPLICATION NUMBER: PCT/US91/08889
FILING DATE: 26-NOV-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: Australian PK3594/90
FILING DATE: 28-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: Australian 88229/91
FILING DATE: 27-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: Baer, Madeline F.
REGISTRATION NUMBER: 36,437
REFERENCE/DOCKET NUMBER: LUD 5244.3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-688-9200
TELEFAX: 212-838-3884
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 1132 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-446-010B-18

Query Match 7.5%; Score 79; DB 1; Length 1132;

Best Local Similarity 24.0%; Pred. No. 4.5;
Matches 48; Conservative 30; Mismatches 64; Indels 58; Gaps 13;

12 ATIACHLDPFVFDGLCRAKFESLFTYDKDTTFQYFKSFKVRINFSNPLSAADARLRL 71
64 AAQACRISP-----LC-----HNLFLYDENTKLMWAPN--RTTVDDKMS-----LRL 105
QY 72 HKTFLQKEMKLYPAQTGHTLHSSHLAPNPDQKFLSPASPVGW--KQVEDATPVNY 129
Db 106 H-----YRMRFT-----NMHGTDNDSQSVRHSPPKQKN-GYEKKKIPDAPLDA 152
QY 130 DLVAISKLGPEKYEHLAATDPTSPVVHVCSQDENESEEE-----ME 175
Db 153 SLELYFAQG---QYDLVYKCLAP-----IRDPKTEODGHDIEHCLGMVLAISHYAMMK 204
QY 176 RMKRPK-PKIIOTRREPTP 194
Db 205 KMQLPELPKDISYKR--YIP 222

RESULT 9

US-08-805-445-18

Sequence 18, Application US/08805445
Patent No. 5821069
GENERAL INFORMATION:
APPLICANT: Wilks, Andrew F.; Ziemiecki, Andrew;
APPLICANT: Harpur, Alisa
TITLE OF INVENTION: No. 5821069el Protein Tyrosine Kinase
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pelte & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
COMPUTER: IBM PS/2

```

; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/805,445
; FILING DATE: 25-FEB-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446,038
; FILING DATE: 19-MAY-1995
; APPLICATION NUMBER: 08/064,067
; FILING DATE: 30-Jun-1993
; APPLICATION NUMBER: PCT/US91/08889
; FILING DATE: 26-No. 5821069-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: Australian PK3594/90
; FILING DATE: 28-No. 5821069-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: Australian 88229/91
; FILING DATE: 27-No. 5821069-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5821069man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5244
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-688-9200
; TELEFAX: 212-838-3884
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1132 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-805-445-18

```

Query Match 7.5%; Score 79; DB 2; Length 1132;

Best Local Similarity 24.0%; Pred. No. 4.5;

Matches 48; Conservative 30; Mismatches 64; Indels 58; Gaps 13;

```

QY 12 ATACHLDPRVFDGLCRKAFESLFTYDKDTTFQYFKSKRVIRNFSNPLSADARLRL 71
DB 64 AAQACRISP-----LC-----HNLFALYDENTKLMYAPN---RTITVDDKMS-----LRL 105
QY 72 HKTEFLGKMKLYPAQTLHIGSSHLAPPNDKQFLISPPASPPVGW--KQVEDATPVINY 129
DB 106 H-----YRMRFYFT-----NWHGTNDNEQSVWRHSPKKQKN-GYEKKKIPDTPLLDA 152
QY 130 DLLYATSKLGPGEKVELHAATDPTSPVVHVHVCSDQNEEEEE-----ME 175
DB 153 SSLEYLFAQG---QYDLVKCLAP-----IRDPKTEQDGHDIENECGLGMAVLAI SHYAMNK 204
QY 176 RMKRPK-PKIIOTRRPEYTP 194
DB 205 KMQLPELPKDISYKR--YIP 222

```

RESULT 10

```

US-08-064-067D-18
; Sequence 18, Application US/08064067D
; Patent No. 5852184
; GENERAL INFORMATION:
; APPLICANT: Wilks, Andrew F.; Ziemiecki, Andrew;
; APPLICANT: Harpur, Ailsa
; TITLE OF INVENTION: No. 5852184el Protein Tyrosine Kinase
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felte & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM PS/2

```

```

; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/064,067D
; FILING DATE: 30-Jun-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/08889
; FILING DATE: 26-No. 5852184-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: Australian PK3594/90
; FILING DATE: 28-No. 5852184-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: Australian 88229/91
; FILING DATE: 27-No. 5852184-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5852184man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5244
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-688-9200
; TELEFAX: 212-838-3884
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1132 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-064-067D-18

```

Query Match 7.5%; Score 79; DB 2; Length 1132;

Best Local Similarity 24.0%; Pred. No. 4.5;

Matches 48; Conservative 30; Mismatches 64; Indels 58; Gaps 13;

```

QY 12 ATACHLDPRVFDGLCRKAFESLFTYDKDTTFQYFKSKRVIRNFSNPLSADARLRL 71
DB 64 AAQACRISP-----LC-----HNLFALYDENTKLMYAPN---RTITVDDKMS-----LRL 105
QY 72 HKTEFLGKMKLYPAQTLHIGSSHLAPPNDKQFLISPPASPPVGW--KQVEDATPVINY 129
DB 106 H-----YRMRFYFT-----NWHGTNDNEQSVWRHSPKKQKN-GYEKKKIPDTPLLDA 152
QY 130 DLLYATSKLGPGEKVELHAATDPTSPVVHVHVCSDQNEEEEE-----ME 175
DB 153 SSLEYLFAQG---QYDLVKCLAP-----IRDPKTEQDGHDIENECGLGMAVLAI SHYAMNK 204
QY 176 RMKRPK-PKIIOTRRPEYTP 194
DB 205 KMQLPELPKDISYKR--YIP 222

```

RESULT 11

```

US-09-066-208-18
; Sequence 18, Application US/09066208
; Patent No. 5910426
; GENERAL INFORMATION:
; APPLICANT: Wilks, Andrew F.; Ziemiecki, Andrew;
; APPLICANT: Harpur, Ailsa
; TITLE OF INVENTION: No. 5910426el Protein Tyrosine Kinase
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felte & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/066,208
; FILING DATE:

```

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/805,445
FILING DATE: 25-FEB-1997
APPLICATION NUMBER: US 08/446,038
FILING DATE: 19-MAY-1995
APPLICATION NUMBER: 08/064,067
FILING DATE: 30-JUN-1993
APPLICATION NUMBER: PCT/US91/08889
FILING DATE: 26-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: Australian PK3594/90
FILING DATE: 28-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: Australian 88229/91
FILING DATE: 27-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5910426man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5244
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-688-9200
TELEFAX: 212-688-3884
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 1132 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-09-066-208-18

Query Match 7.5%; Score 79; DB 2; Length 1132;
Best Local Similarity 24.0%; Pred. No. 4.5;
Matches 48; Conservative 30; Mismatches 64; Indels 58; Gaps 13;

QY 12 ATACHDPRVFDGLCRAFESIFRTYDQDTFOYKSKFRVINSNPLSADARLR 71
DB 64 AAOACRISP-----LC-----HNLFALYDENTKLTWAPN--RTTYDDKMS-----LRL 105
QY 72 HKTEFLGKEMKLYPAQTLHGSSHLAPPNDKQFLISPPASPPGV--KQVEDATPVINY 129
DB 106 H-----YRKRFYFT-----NMHGTDNDSQSVWRHSPKQKN-GYEKKKIIPATPLLDA 152
QY 130 DLVAISKLGEGEYELHAATDPTPSVVHVHCSDDQNEEBEEB-----ME 175
DB 153 SSLEYLFAQG---QYDLVKCLAP-----IRDPKTEODGHDIENECIGMAVLAISHVAMK 204
QY 176 RMKRPK-PKIQTRRPETP 194
DB 205 KMQLPELPKDISYKR--YIP 222

RESULT 12
US-08-097-997A-11
Sequence 11, Application US/08097997A
Patent No. 5728536
GENERAL INFORMATION:
APPLICANT: Ihle, James N.
APPLICANT: Silvenoinen, Olli
APPLICANT: Mituhnu, Bruce A.
APPLICANT: Ouellet, Frederick W.
TITLE OF INVENTION: Jak kinases and Regulation of Cytokine Signal
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/097,997A
FILING DATE: 29-JULY-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fox, Samuel L.
REGISTRATION NUMBER: 30,353
REFERENCE/DOCKET NUMBER: 0656.0370000/SLF/GKT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1142 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-097-997A-11

Query Match 7.5%; Score 79; DB 1; Length 1142;
Best Local Similarity 24.0%; Pred. No. 4.5;
Matches 48; Conservative 30; Mismatches 64; Indels 58; Gaps 13;

QY 12 ATACHDPRVFDGLCRAFESIFRTYDQDTFOYKSKFRVINSNPLSADARLR 71
DB 52 AAOACRISP-----LC-----HNLFALYDENTKLTWAPN--RTTYDDKMS-----LRL 93
QY 72 HKTEFLGKEMKLYPAQTLHGSSHLAPPNDKQFLISPPASPPGV--KQVEDATPVINY 129
DB 94 H-----YRKRFYFT-----NMHGTDNDSQSVWRHSPKQKN-GYEKKKIIPATPLLDA 140
QY 130 DLVAISKLGEGEYELHAATDPTPSVVHVHCSDDQNEEBEEB-----ME 175
DB 141 SSLEYLFAQG---QYDLVKCLAP-----IRDPKTEODGHDIENECIGMAVLAISHVAMK 192
QY 176 RMKRPK-PKIQTRRPETP 194
DB 193 KMQLPELPKDISYKR--YIP 210

RESULT 13
US-08-665-574C-11
Sequence 11, Application US/08665574C
Patent No. 6136595
GENERAL INFORMATION:
APPLICANT: Ihle, James N.
APPLICANT: Silvenoinen, Olli
APPLICANT: Mituhnu, Bruce A.
TITLE OF INVENTION: Jak kinases and Regulation of Cytokine
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/665,574C
FILING DATE: 18-JUN-1996
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/282,012
FILING DATE: 29-JUL-1994
PRIOR APPLICATION DATA:

```

; APPLICATION NUMBER: 08/097,997
; FILING DATE: 29-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/118,968
; FILING DATE: 09-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Bugaieky, Lawrence B.
; REGISTRATION NUMBER: 35,086
; REFERENCE/DOCKET NUMBER: 0656.0370002/SLF/LBB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1142 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-665-574C-11

```

```

Query Match          7.5% Score 79; DB 4; Length 1142;
Best Local Similarity 24.0%; Pred. No. 4.5;
Matches 48; Conservative 30; Mismatches 64; Indels 58; Gaps 13;

QY 12 ATACHLDPRVVDGLCRKAFESLRTYDKDTTFQYFKSKVRINFSNPLSAADARLRL 71
DB 52 AAQACRISP-----LC-----HNLFALYDENTKLYAPN---RTITVDDKMS-----LRL 93
QY 72 HKTEFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPPVGW--KQVEDATPVINY 129
DB 94 H-----YRMRFYFT-----NWHGTNDNEQSVWRHSPKQKN-GYEKKKIPDATPLLDA 140
QY 130 DLLVAISKLGPGKGYELHAATDPTPSVVHVCSDEQNEEEEE-----ME 175
DB 141 SSLEYLFAQG---QYDLVKCLAP-----IRDPKTEQDGHDIENECGLMAVLAISHYAMWK 192
QY 176 RMKRPK-PKIIOTRPEYTP 194
DB 193 KMQLPELPKDISYKR--YIP 210

```

```

RESULT 14
US-08-946-994-11
; Sequence 11, Application US/08946994
; Patent No. 6210654
; GENERAL INFORMATION:
; APPLICANT: Ihle, James N.
; APPLICANT: Silvenoinen, Ollie
; APPLICANT: Withuhn, Bruce A.
; APPLICANT: Quelle, Frederick W.
; TITLE OF INVENTION: Jak Kinases and Regulation of Cytokine Signal
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/946,994
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/665,574
; FILING DATE: 18-JUN-1996
; REGISTRATION NUMBER: 08/282,012

```

```

; FILING DATE: 29-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/097,997
; FILING DATE: 29-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/118,968
; FILING DATE: 09-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Fox, Samuel L.
; REGISTRATION NUMBER: 30,353
; REFERENCE/DOCKET NUMBER: 0656.0370002/SLF/GKT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1142 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-946-994-11

```

```

Query Match          7.5% Score 79; DB 4; Length 1142;
Best Local Similarity 24.0%; Pred. No. 4.5;
Matches 48; Conservative 30; Mismatches 64; Indels 58; Gaps 13;

QY 12 ATACHLDPRVVDGLCRKAFESLRTYDKDTTFQYFKSKVRINFSNPLSAADARLRL 71
DB 52 AAQACRISP-----LC-----HNLFALYDENTKLYAPN---RTITVDDKMS-----LRL 93
QY 72 HKTEFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPPVGW--KQVEDATPVINY 129
DB 94 H-----YRMRFYFT-----NWHGTNDNEQSVWRHSPKQKN-GYEKKKIPDATPLLDA 140
QY 130 DLLVAISKLGPGKGYELHAATDPTPSVVHVCSDEQNEEEEE-----ME 175
DB 141 SSLEYLFAQG---QYDLVKCLAP-----IRDPKTEQDGHDIENECGLMAVLAISHYAMWK 192
QY 176 RMKRPK-PKIIOTRPEYTP 194
DB 193 KMQLPELPKDISYKR--YIP 210

```

```

RESULT 15
US-08-357-598-7
; Sequence 7, Application US/08357598
; Patent No. 5705625
; GENERAL INFORMATION:
; APPLICANT: Civin, Curt I.
; APPLICANT: Small, Donald
; TITLE OF INVENTION: NOVEL PROTEIN TYROSINE KINASE, JAK3
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/357,598
; FILING DATE: 15-DEC-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/033001
; TELECOMMUNICATION INFORMATION:

```

Sun Dec 15 08:38:08 2002

us-09-782-953-3.ral

TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1154 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-357-598-7

Query Match 7.5%; Score 79; DB 1; Length 1154;
Best Local Similarity 24.0%; Pred. No. 4.6; Mismatches 64; Indels 58; Gaps 13;
Matches 48; Conservative 30;

QY 12 ATACHLDPRVFDGLGCRAKESLFRITDKDTTFQYFKSFKVRINFNSNPLSADARLRL 71
DB 64 AAQCRISP-----LC---HNLFALYDENTKLMYAPN--RTITVDKMS-----LRL 105
QY 72 HKTFLGKEMKLYFAQTUHGSSHLAPNDKQFLISPPASPPVGW--KQVEDATPVINY 129
DB 106 H-----YRMRFYFT-----NMHGTDNDSQSVWRHSPKKQKN-GYEKKKIPDATPLIDA 152
QY 130 DLIVAIKLGGEKYLHAATDPTPSVVVHVCSDSQENESEEE-----ME 175
DB 153 SLEYLFAQG---QYDLVKCLAP-----IRDPKTEQDGHDIENECLEMAVLAISHYAMMK 204
QY 176 RMKRPR-PKIIQTRRPEYTP 194
DB 205 KMQLPELPKDISYKR--YIP 222

Search completed: December 11, 2002, 11:39:12
Job time: 15.3524 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 14, 2002, 21:32:16 ; Search time 33.0556 Seconds
(without alignments)
2368.771 Million cell updates/sec

Title: US-09-782-953-3

Perfect score: 1048

Sequence: 1 MEEVDQLPSATACHLDP.....RPKIKIOTRRPEYTPHLS 198

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 355320 seqs, 197730502 residues

Total number of hits satisfying chosen parameters: 710640

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09782953/runat_11122002_114431_17465/app.query.fasta_1.1173
-DB=Published Applications NA -OFMT=fastcap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09782953@cgn 1 1 38 @runat_11122002_114431_17465
-NCFU=6 -ICFU=3 -NO XLPXY -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELEXT=7

Database :

Published Applications NA:

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	Query Match	Length	ID	Description
1	100.0	597	10	US-09-782-953-2	Sequence 2, Appli
2	95.7	2212	10	US-09-782-953-14	Sequence 14, Appl
3	2358	10	US-09-925-302-347	Sequence 347, App	Sequence 5, Appli
4	597	10	US-09-782-953-5		

5	872.5	83.3	2331	10	US-09-782-953-11	Sequence 11, Appl
6	872.5	83.3	2355	12	US-10-044-090-255	Sequence 255, App
7	847.5	80.9	2173	10	US-09-880-107-3340	Sequence 3340, Ap
8	749.5	71.5	599	10	US-09-782-953-1	Sequence 1, Appli
9	613.5	58.5	594	10	US-09-782-953-8	Sequence 8, Appli
10	596.5	56.9	3184	10	US-09-954-456-497	Sequence 497, App
11	596.5	56.9	3184	10	US-09-782-953-17	Sequence 17, Appl
12	596	56.9	828	10	US-09-782-953-20	Sequence 20, Appl
13	595	56.8	720	10	US-09-782-953-23	Sequence 23, Appl
14	285.5	27.2	412	10	US-09-864-761-10388	Sequence 10388, A
15	285.5	27.2	446	10	US-09-864-761-2064	Sequence 2064, Ap
16	266.5	25.4	486	10	US-09-864-761-809	Sequence 809, App
17	200	19.1	365	10	US-09-728-445-736	Sequence 736, App
18	187.5	17.9	123	10	US-09-864-761-17592	Sequence 17592, A
19	181	17.3	111	10	US-09-864-761-27019	Sequence 27019, A
20	140	13.4	85	10	US-09-864-761-18808	Sequence 18808, A
21	90.5	8.6	12308	9	US-09-854-133-422	Sequence 422, App
22	90.5	8.6	12308	10	US-09-738-973-422	Sequence 422, App
23	87	8.3	4600	9	US-09-736-457-1797	Sequence 1797, Ap
24	87	8.3	4600	9	US-09-902-941-1797	Sequence 1797, Ap
25	83.5	8.0	32191	10	US-09-764-877-3374	Sequence 3374, Ap
26	82.5	7.9	530	10	US-09-864-761-15897	Sequence 15897, A
27	81.5	7.8	2095	12	US-10-044-090-665	Sequence 665, App
28	81	7.7	3138	10	US-09-880-107-1716	Sequence 1716, App
29	80.5	7.7	1211	9	US-09-981-876-123	Sequence 123, App
30	80.5	7.7	6799	9	US-09-902-941-1883	Sequence 1883, Ap
31	80.5	7.7	16798	10	US-09-954-456-594	Sequence 594, App
32	80	7.6	552	10	US-09-771-161A-29	Sequence 29, Appl
33	79	7.5	2027	10	US-09-880-107-2379	Sequence 2379, Ap
34	79	7.5	3541	10	US-09-954-456-2210	Sequence 2210, Ap
35	79	7.5	3775	10	US-09-964-469-3	Sequence 3, Appli
36	79	7.5	3775	10	US-09-964-469-3	Sequence 301, App
37	79	7.5	36651	10	US-09-974-300-1099	Sequence 1099, Ap
38	78.5	7.5	569	9	US-09-883-096-1	Sequence 1, Appli
39	78.5	7.5	738	10	US-09-349-755-33	Sequence 33, Appl
40	78.5	7.5	4108	10	US-09-166-334-33	Sequence 33, Appl
41	77	7.3	1086	9	US-09-350-206-33	Sequence 94, Appl
42	77	7.3	1086	9	US-09-349-755-33	Sequence 31, Appl
43	77	7.3	1086	10	US-09-712-363-94	
44	77	7.3	1440	9	US-09-349-755-31	
45	77	7.3	2218	9		

ALIGNMENTS

RESULT 1

US-09-782-953-2
; Sequence 2, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: US/09/782,953
; CURRENT APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 597
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(594)
US-09-782-953-2

Alignment Scores:
Pred. No.: 2.07e-119
Score: 1048.00
Length: 597
Matches: 198

Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 0

US-09-782-953-3 (1-198) x US-09-782-953-2 (1-597)

QY 1 MetGLUGLVALAspLeuGlnAspLeuProSerAlaThrIleAlaCysHISLeuAspPro 20
 DB 1 ATGAGAGAGGTGATCTGAGAGACCTGCCAGCCAGCATGCTGCTGCACCTGGACCCG 60
 QY 21 ArgValPheValAspIleuCysArgAlaIlyPheGluSerLeuPheArgThrTyArg 40
 DB 61 CGCGTGTCTGAGCGGCTGCTGCCGCCAATTGAACTCCCTTCAGAACATATGAC 120
 QY 41 LysAspThrThrPheGlnTyRheIySerPheIyAspValArgIleAspPheSerAsn 60
 DB 121 AAGGACACCACTTCCAGTATTTAAAGCTTCAACGTCGCGATTAACCTTCAGCAAC 180
 QY 61 ProIleSerAlaIleAspAlaArgLeuArgLeuHISlyeThrGluPheLeuGlyIlyGlu 80
 DB 181 CCTTATCTGCAGCCGATGCCAGGCTGCCGCTGCACAAAGACGAGTTCCTGGGAAAGAA 240
 QY 81 MetIyLeuTyRheAlaGlnThrLeuHISIleGlySerSerHISLeuAlaProProAsn 100
 DB 241 ATGAAGTTATTTTGGCTCAGACTTACACATAGAAAGTTCACCTGGCTCCGCCAAT 300
 QY 101 ProAspIyGlnPheLeuIleSerProAlaSerProProValGlyTyIlyGlnVal 120
 DB 301 CCCGACAAAGTTCCTCATCTCCCTCCGCTCCCTCCGCTCCGCTGGGAAACAAAGTA 360
 QY 121 GluAspAlaThrProValIleAsnTyRAspLeuLeuTyRAlIleSerIyIleGlyPro 140
 DB 361 GAAGATGCCACCCCGCATTAATTCATCTTTATATGCACTTCGAAAGCTGGGCCA 420
 QY 141 GlyIyIyTyRGIyLeuHISAlaIleThrAspProThrProSerValIleValHISVal 160
 DB 421 GGAGAGAAATAGAACTGATGACGACGACCAACCCCACTCCCAAGTGGTGCACGTG 480
 QY 161 CysGluSerAspGlnIyAsnGluGluGluGluIyMetGluIyGlyMetIyAspPro 180
 DB 481 TGTGAGAGTACCAAGAAATGAGAGAAAGAAAGATGAGAAATGAAGAACCC 540
 QY 181 LysProIyIleIleGlnThrArgArgProGluTyRThrProIleHISLeuSer 198
 DB 541 AAGCCCAAAATATCCAGACGAGACCGAGTACACCGATCCACTTACG 594

RESULT 2

US-09-782-953-14

Sequence 14, Application US/09782953
 Patent No. US20020150953A1
 GENERAL INFORMATION:
 APPLICANT: WILLIAMS, R. SANDERS
 APPLICANT: ROTHERMEL, BEVERLY
 TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
 TITLE OF INVENTION: CALCIUM INTERACTING PROTEIN (MCIP)
 FILE REFERENCE: UTSD-674P21
 CURRENT APPLICATION NUMBER: US/09/782,953
 PRIOR FILING DATE: 2001-02-13
 PRIOR APPLICATION NUMBER: 60/216,601
 PRIOR FILING DATE: 2000-07-07
 NUMBER OF SEQ ID NOS: 27
 SOFTWARE: Patent Ver. 2.1
 SEQ ID NO 14
 LENGTH: 2212
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (25)..(615)
 US-09-782-953-14

Alignment Scores:

Pred. No.: 5.26e-113 Length: 2212
 Score: 1002.50 Matches: 191
 Percent Similarity: 97.98% Conservative: 3
 Best Local Similarity: 96.46% Mismatches: 3
 Query Match: 95.66% Indels: 1
 Gaps: 1

US-09-782-953-3 (1-198) x US-09-782-953-14 (1-2212)

QY 1 MetGLUGLVALAspLeuGlnAspLeuProSerAlaThrIleAlaCysHISLeuAspPro 20
 DB 25 ATGAGAGAGGTGATCTGAGAGACCTGCCAGCCAGCATGCTGCTGCACCTGGACCCG 84
 QY 21 ArgValPheValAspIleuCysArgAlaIlyPheGluSerLeuPheArgThrTyArg 40
 DB 85 CGCGTGTCTGAGCGGCTGCTGCCGCCAATTGAACTCCCTTCAGAACATATGAC 144
 QY 41 LysAspThrThrPheGlnTyRheIySerPheIyAspValArgIleAspPheSerAsn 60
 DB 145 AAGGACATCACTTCCAGTATTTAAAGCTTCAACGTCGCGATTAACCTTCAGCAAC 204
 QY 61 ProIleSerAlaIleAspAlaArgLeuArgLeuHISlyeThrGluPheLeuGlyIlyGlu 80
 DB 205 CCTTATCTGCAGCCGATGCCAGGCTGCCGCTGCACAAAGACGAGTTCCTGGGAAAGAA 264
 QY 81 MetIyLeuTyRheAlaGlnThrLeuHISIleGlySerSerHISLeuAlaProProAsn 100
 DB 265 ATGAAGTTATTTTGGCTCAGACTTACACATAGAAAGTTCACCTGGCTCCGCCAAT 324
 QY 101 ProAspIyGlnPheLeuIleSerProAlaSerProProValGlyTyIlyGlnVal 120
 DB 325 CGAGACAAAGTTCCTCATCTCCCTCCGCTCCCTCCGCTCCGCTGGGAAACAAAGTG 384
 QY 121 GluAspAlaThrProValIleAsnTyRAspLeuLeuTyRAlIleSerIyIleGlyPro 140
 DB 385 GAAGATGCCACCCCGCATTAATTCATCTTTATATGCACTTCGAAAGCTGGGCCA 444
 QY 141 GlyIyIyTyRGIyLeuHISAlaIleThrAspProThrProSerValIleValHISVal 160
 DB 445 GGGGAAAGTATGAATTCACGACGACACTACACCACTCCCAAGTGGTGCACGTG 504
 QY 161 CysGluSerAspGlnIyAsnGluGluGluGluIyMetGluIyGlyMetIyAspPro 180
 DB 505 TGTGAGAGTACCAAGAG--AAGAGGAAAGAAAGAAAGATGAGAAATGAAGAACCT 561
 QY 181 LysProIyIleIleGlnThrArgArgProGluTyRThrProIleHISLeuSer 198
 DB 562 AAGCCCAAAATATCCAGACGAGGCGCGAGTACACCGGATCCACTTACG 615

RESULT 3

US-09-925-302-347

Sequence 347, Application US/09925302
 Patent No. US2002004941A1
 GENERAL INFORMATION:
 APPLICANT: ROSEN ET AL.
 TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
 FILE REFERENCE: PA104
 CURRENT APPLICATION NUMBER: US/09/925,302
 PRIOR FILING DATE: 2001-08-10
 PRIOR APPLICATION NUMBER: PCT/US00/05918
 PRIOR FILING DATE: 2000-03-08
 PRIOR APPLICATION NUMBER: 60/124,270
 PRIOR FILING DATE: 1999-03-12
 NUMBER OF SEQ ID NOS: 896
 SOFTWARE: Patent Ver. 2.0
 SEQ ID NO 347
 LENGTH: 2358
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-09-925-302-347

Alignment Scores:

Pred. No.: 1.86e-107 Length: 2358

Score: 957.50 Matches: 188
Percent Similarity: 96.46% Conservative: 3
Best Local Similarity: 94.95% Mismatches: 6
Query Match: 91.36% Indels: 2
DB: 10 Gaps: 1

US-09-782-953-3 (1-198) x US-09-925-302-347 (1-2358)

Qy 1 MetGluValAspLeuGlnAspLeuProSerAlaThrIleAlaCysHisLeuAspPro 20
Db 114 ATGGAGAGGTGGAGCTGGAGACCTGCCAGCGCCACCATCGCTGTCACCTGGACCG 173
Qy 21 ArgValPheValAspGlyLeuCysArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40
Db 174 CGCGTTCGTGGACGGCTGTGCCGGGCAAAATTTAGTCCCTCTTTAGGACGATGAC 233
Qy 41 LysAspThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60
Db 234 AAGGACATCACCTTTTCAGTATTTTAAGAGCTTCAAAACGAGTCAGAAATAAATTCAGCAAC 293
Qy 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
Db 294 CCCTTCTCCGACGACATGCCAGCTCCAGCTGCATAAGACTGAGTTTCTGGGAAAGGAA 353
Qy 81 MetLysLeuTyrPheAlaGlnThrLeuHisLysThrGluPheLeuGlyLysGlu 100
Db 354 ATGAAGTTATATTTGCTCAGACCTTACACATAGGAAGCTCACACCTGGCTCCGCA-AAAT 412
Qy 101 ProAspLysGlnPheLeuSerProAlaSerProValGlyTyrLysGlnVal 120
Db 413 CCAGAACAGCAGTTCTGATCTCCCTCCGCTCTCCGSCAGTGGAGTGAACAAAGTG 472
Qy 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro 140
Db 473 GAAGATGCGACCCAGTCATAACTATGATCTCTTATATGCCATCTCCAGCTGGGGCCA 532
Qy 141 GlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160
Db 533 GGGGAAAGTATGAATTCACGCGAGCTGACACCACTCCAGCTGGTGGTCCATGTA 592
Qy 161 CysGluSerAspGlnGluAsnGluGluGluGluMetGluArgMetLysArgPro 180
Db 593 TGTAGAGTGCATCAAGAG---AAGGAGGAAGAGAGAAATGGAAGATGAGGAGACCT 649
Qy 181 LysProLysIleIleGlnThrArgArgProGluTyrThrProIleHisLeuSer 198
Db 650 AAGCCAAAATTTATCCAGACGAGGAGGCGGAGTACAGCGGATCCACCTCAGC 703

RESULT 4

US-09-782-953-5
; Sequence 5, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)
; FILE REFERENCE: UTSD:674P21
; CURRENT APPLICATION NUMBER: US/09/782,953
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 597
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(594)
US-09-782-953-5

Alignment Scores:
Pred. No.: 8,33e-104 Length: 597
Score: 920.50 Matches: 178
Percent Similarity: 92.42% Conservative: 5
Best Local Similarity: 89.90% Mismatches: 12
Query Match: 87.83% Indels: 3
DB: 10 Gaps: 1

US-09-782-953-3 (1-198) x US-09-782-953-5 (1-597)

Qy 4 ValAspLeuGlnAspLeuPro-----SerAlaThrIleAlaCysHisLeuAspPro 20
Db 1 ATGATATTTAGGACTTTAGCTACAAATTTTAGCTCCCTGATTGCTTGTGTGCAACACGAT 60
Qy 21 ArgValPheValAspGlyLeuCysArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40
Db 61 GATGCTTTCAGCGAAAGTGAGACCGAGGCGCAAAATTTGAATCCCTCTTCAGAACATATGAC 120
Qy 41 LysAspThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60
Db 121 AAGGACACCATCTTCCAGTATTTTAAGAGCTTCAACGTGCCGATTAACCTTCAGCAAC 180
Qy 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
Db 181 CCTTATCTGACGCGATGCCAGCTCGCGCTGCACAGACCGAGTTCTTGGGAAAGGAA 240
Qy 81 MetLysLeuTyrPheAlaGlnThrLeuHisLysThrGluPheLeuGlyLysGlu 100
Db 241 ATGAAGTTGATTTTGCCTCAGACTTTACACATAGGAAGTTTCACACCTGGCTCCGCCCAAT 300
Qy 101 ProAspLysGlnPheLeuSerProAlaSerProValGlyTyrLysGlnVal 120
Db 301 CCCGACAAACAGTTCCTCATCTCCCTCCGCTCTCTCCGTTGGCTGGAACAAAGTA 360
Qy 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro 140
Db 361 GAAGATGCGACCCCGCTCATAAATTCAGTCTTTTATATGCCATCTCCAGCTGGGGCCA 420
Qy 141 GlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160
Db 421 GGAGAGAAAGTATGAATTCATGTCAGCGACAGACCCCACTCCAGCTGGTGGTCCACGCTG 480
Qy 161 CysGluSerAspGlnGluAsnGluGluGluGluMetGluArgMetLysArgPro 180
Db 481 TGTAGAGTGCACCAAGAAATGAGGAGGAAGAGAGATGGAGAGATGAGAGATGAGAGACCC 540
Qy 181 LysProLysIleIleGlnThrArgArgProGluTyrThrProIleHisLeuSer 198
Db 541 AAGCCAAAATTTATCCAGACGAGGAGCGGAGTACACCGGATCCACCTTAGC 594

RESULT 5

US-09-782-953-11
; Sequence 11, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)
; FILE REFERENCE: UTSD:674P21
; CURRENT APPLICATION NUMBER: US/09/782,953
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 2331
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (144)...(734)
;

US-09-782-953-11

Alignment Scores:

Pred. No.:	4,61e-97	Length:	2331
Score:	872.50	Matches:	168
Percent Similarity:	93.09%	Conservative:	7
Best Local Similarity:	89.36%	Mismatches:	12
Query Match:	83.25%	Indels:	1
DB:	10	Gaps:	1

US-09-782-953-3 (1-198) x US-09-782-953-11 (1-2331)

```

QY 11 SerAlaThrIleAlaCyHisIleuAspProArgValPheValAspGlyLeuCyArgAla 30
   |||
Db 174 AGTCCCTGATTCCTGCTGGCAACAGTATCTTCAGCGAAAGTGAACCGAGGCC 233
   |||

QY 31 LysPheGluSerLeuPheArgThrTyrAspLysAspThrThrpheGlnTyrPheLysSer 50
   |||
Db 234 AAATTGAGTCCCTCTTAGAGCATATGACAGACATCACCTTTCAGTATTTTAAAGAC 293
   |||

QY 51 PheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAspAlaArgLeuArg 70
   |||
Db 294 TTCAAACGAGTCAGAAATAAATTCTCAGCAACCCCTTCTCCGAGCAGATGCCAGCTCAG 353
   |||

QY 71 LeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyrPheAlaGlnThrLeuHis 90
   |||
Db 354 CTGCATTAAGACTGAGTTCTGGCAAAAGAAATGAAGTAAATTTTGTCTCAGACCTTACAC 413
   |||

QY 91 IleGlySerSerHisLeuAlaProProAsnProAspLysGlnPheLeuIleSerProPro 110
   |||
Db 414 ATGGAAGCTCACACCTGGCTCCGCCAAATCCAGACAGACAGTTTCTGATCTCCCTCCC 473
   |||

QY 111 AlaSerProProValGlyTyrLysGlnValGluAspAlaThrProValIleAsnTyrAsp 130
   |||
Db 474 GCCTCTCCGCGAGTGGAGTGAACAAGTGAAGATCGACCCAGCATTAACATATGAT 533
   |||

QY 131 LeuLeuTyrAlaIleSerLysLeuGlyProGlyGluLysTyrGluLeuHisAlaAlaThr 150
   |||
Db 534 CTTTATATGCGCATCTCCAAAGCTGGGCCAGGGGAAAGTAAATTTGACGACGAGACT 593
   |||

QY 151 AspProThrProSerValValIleHisValCysGluSerAspGlnGluAsnGluGluGlu 170
   |||
Db 594 GACACCACTCCACGCGGTGGTGCATGTATGAGATATCAAGAG--AAGAGAGAA 650
   |||

QY 171 GluGluGluMetLysArgProLysProLysIleIleGlnThrArgArgPro 190
   |||
Db 651 GAAGAGGAAATGAAAGAAATGAGAGACCTTAAGCCAAAATTTATCCAGACAGAGAGCCG 710
   |||

QY 191 GluTyrThrProIleHisLeuSer 198
   |||
Db 711 GAGTACACGCCGATCCACCTCAGC 734
   |||

```

RESULT 6

US-10-044-090-255

Sequence 255, Application US/10044090

Patent No. US20020137081A1

GENERAL INFORMATION:

APPLICANT: Olga Bandman

TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION

FILE REFERENCE: PA-0028 US

CURRENT APPLICATION NUMBER: US/10/044,090

CURRENT FILING DATE: 2002-01-09

NUMBER OF SEQ ID NOS: 850

SOFTWARE: PERL Program

SEQ ID NO 255

LENGTH: 2355

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc_feature

OTHER INFORMATION: Incyte ID No. US20020137081A1 551189CB1

US-10-044-090-255

Alignment Scores:

Pred. No.:	4,68e-97	Length:	2355
Score:	872.50	Matches:	168
Percent Similarity:	93.09%	Conservative:	7
Best Local Similarity:	89.36%	Mismatches:	12
Query Match:	83.25%	Indels:	1
DB:	12	Gaps:	1

US-09-782-953-3 (1-198) x US-10-044-090-255 (1-2355)

```

QY 11 SerAlaThrIleAlaCyHisIleuAspProArgValPheValAspGlyLeuCyArgAla 30
   |||
Db 181 AGTCCCTGATTCCTGCTGGCAACAGTATCTTCAGCGAAAGTGAACCGAGGCC 240
   |||

QY 31 LysPheGluSerLeuPheArgThrTyrAspLysAspThrThrpheGlnTyrPheLysSer 50
   |||
Db 241 AAATTGAGTCCCTCTTAGAGCATATGACAGACATCACCTTTCAGTATTTTAAAGAC 300
   |||

QY 51 PheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAspAlaArgLeuArg 70
   |||
Db 301 TTCAAACGAGTCAGAAATAAATTCTCAGCAACCCCTTCTCCGAGCAGATGCCAGCTCAG 360
   |||

QY 71 LeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyrPheAlaGlnThrLeuHis 90
   |||
Db 361 CTGCATTAAGACTGAGTTCTGGCAAAAGAAATGAAGTAAATTTTGTCTCAGACCTTACAC 420
   |||

QY 91 IleGlySerSerHisLeuAlaProProAsnProAspLysGlnPheLeuIleSerProPro 110
   |||
Db 421 ATGGAAGCTCACACCTGGCTCCGCCAAATCCAGACAGACAGTTTCTGATCTCCCTCCC 480
   |||

QY 111 AlaSerProProValGlyTyrLysGlnValGluAspAlaThrProValIleAsnTyrAsp 130
   |||
Db 481 GCCTCTCCGCGAGTGGAGTGAACAAGTGAAGATCGACCCAGCATTAACATATGAT 540
   |||

QY 131 LeuLeuTyrAlaIleSerLysLeuGlyProGlyGluLysTyrGluLeuHisAlaAlaThr 150
   |||
Db 541 CTTTATATGCGCATCTCCAAAGCTGGGCCAGGGGAAAGTAAATTTGATTCACGACGAGACT 600
   |||

QY 151 AspProThrProSerValValIleHisValCysGluSerAspGlnGluAsnGluGluGlu 170
   |||
Db 601 GACACCACTCCACGCGGTGGTGCATGTATGAGATATCAAGAG--AAGAGAGAA 657
   |||

QY 171 GluGluGluMetLysArgProLysProLysIleIleGlnThrArgArgPro 190
   |||
Db 658 GAAGAGGAAATGAAAGAAATGAGAGACCTTAAGCCAAAATTTATCCAGACAGAGAGCCG 717
   |||

QY 191 GluTyrThrProIleHisLeuSer 198
   |||
Db 718 GAGTACACGCCGATCCACCTCAGC 741
   |||

```

RESULT 7

US-09-880-107-3340

Sequence 3340, Application US/09880107

Patent No. US20020142981A1

GENERAL INFORMATION:

APPLICANT: Horne, Darci T.

APPLICANT: Vockley, Joseph G.

APPLICANT: Scherf, Uwe

APPLICANT: Gene Logic, Inc.

TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer

FILE REFERENCE: 44921-5028-WO

CURRENT APPLICATION NUMBER: US/09/880,107

CURRENT FILING DATE: 2001-06-14

PRIOR APPLICATION NUMBER: US 60/211,379

PRIOR FILING DATE: 2000-06-14

PRIOR APPLICATION NUMBER: US 60/237,054

PRIOR FILING DATE: 2000-10-02

NUMBER OF SEQ ID NOS: 3950

SOFTWARE: Patent In Ver. 2.1

SEQ ID NO 3340

LENGTH: 2173

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc_feature

OTHER INFORMATION: Incyte ID No. US20020137081A1 551189CB1

US-10-044-090-255

FEATURE:
OTHER INFORMATION: Genbank Accession No. US20020142981A1 U28833
US-09-880-107-3340

Alignment Scores:
Pred. No.: 4,77e-94 Length: 2173
Score: 847.50 Matches: 162
Percent Similarity: 97.63% Conservative: 3
Best Local Similarity: 95.86% Mismatches: 3
Query Match: 80.87% Indels: 1
DB: 10 Gaps: 1

US-09-782-953-3 (1-198) x US-09-880-107-3340 (1-2173)

QY 30 AlaLysPheGluSerLeuPheArgThrTyrAspLysAspThrThrPheGlnTyrPheLys 49
DB 58 GCCAAATTTGAGTCCCTCTTTAGGACGATGACAAGGACATCACCTTTCAGTATTTAAG 117
QY 50 SerPheLysAlaGValAlaGleAsnPheSerAsnProLeuSerAlaAlaAspAlaArgLeu 69
DB 118 AGCTTCAACAGAGTCAGAAATAAATTCAGCAACCCCTTCTCCGAGCAGATGCCAGGCTC 177
QY 70 ArgLeuHisLysThrGluPheLeuGluLysGluMetLysLeuTyrPheAlaGlnThrLeu 89
DB 178 CAGCTGCATAAGACTGAGTTTCTGGAAAGGAATGAGTTATATTTTCTCAGACCTTA 237
QY 90 HisLeuGlySerSerHisLeuAlaProProAsnProAspLysGlnPheLeuLysSerPro 109
DB 238 CACATAGGAAGCTCACACCTGGCTCCGCCAAATCCAGCAAGCAGTTTCTGATCTCCCT 297
QY 110 ProAlaSerProProValGlyTrpLysGlnValGluAspAlaThrProValLysAsnTyr 129
DB 298 CCCGCTCTCCGCGAGTGGATGGAACAAGTGGAAATGCGACCCAGTCATAAACTAT 357
QY 130 AspLeuLeuTyrAlaLysSerLysLeuGlyProGlyGluLysTyrGluLeuHisAlaAla 149
DB 358 GATCTCTATATGATCCATCTCCAACTCGGCCAGGGGAAAGATGATGAAATGCGACGCG 417
QY 150 ThrAspProThrProSerValValHisValCysGluSerAspGlnGluAsnGluGlu 169
DB 418 ACTGACACCACTCCACCGTGGTGTCCATGTATGTGAGAGTGATCAAGAG---AAGGAG 474
QY 170 GluGluGluGluMetGluArgMetLysArgProLysProLysLeuLysLeuGlnThrArg 189
DB 475 GAAGAAGGAGGAATGGAAAGATGAGGAGACCTTAAGCAAAAATTTATCCAGACGAGG 534
QY 190 ProGluTyrThrProLeuHisLeuSer 198
DB 535 CCGAGTACACGCGGATCCACCTCAGC 561

RESULT 8

US-09-782-953-1
; Sequence 1, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)
; FILE REFERENCE: UTSD:674P21
; CURRENT APPLICATION NUMBER: US/09/782,953
; PRIOR FILING DATE: 2001-02-13
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-782-953-1

Alignment Scores:

Pred. No.: 3.07e-66 Length: 594
Score: 613.50 Matches: 119

Pred. No.: 7.07e-83 Length: 599
Score: 749.50 Matches: 151
Percent Similarity: 90.12% Conservative: 4
Best Local Similarity: 87.79% Mismatches: 14
Query Match: 71.52% Indels: 4
DB: 10 Gaps: 1
US-09-782-953-3 (1-198) x US-09-782-953-1 (1-599)
QY 5 AspLeuGlnAspLeuPro-----SerAlaThrIleAlaCysHisLeuAspProArg 21
DB 83 GATTTTGGGACTTTAGCTACAATTTAGCTCCCTGATTGTTGTGGCAACGATGAT 142
QY 22 ValPheValAspGlyLeuCysArgAlaLysPheGluSerLeuPheArgThrTyrAspLys 41
DB 143 GTCTTCAGGGAAGAGTGAGACCGAGGCAAAATTTGAATCCCTCTTCAGAACATATGACAAG 202
QY 42 AspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsnPro 61
DB 203 GACACCACCTTCCAGTATTTTAAAGAGCTTCAACGTTCCGATATAACTTCAGCAACCCC 262
QY 62 LeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGluMet 81
DB 263 TTATCTGCAGCGGATGCCAGGCTGGCGGTGCAAGAGCCGAGTTCTCTGGGAAGGAAATG 322
QY 82 LysLeuTyrPheAlaGlnThrLeuHisLysGlySerSerHisLeuAlaProProAsnPro 101
DB 323 AGTTGTATTTTGTCTCAGACTTTACATAGGAAGTTTACACCTCGCTCCGCAAT-CCC 381
QY 102 AspLysGlnPheLeuLysSerProProAlaSerProProValGlyTrpLysGlnValGlu 121
DB 382 GACAAACAGTCTCTCATCTCCCTCCGCTCTCTCCGCTGCTGGTGGTGGAAACAGTAGAA 441
QY 122 AspAlaThrProValLysLeuAsnTyrAspLeuLeuTyrAlaLysSerLysLeuGlyProGly 141
DB 442 GATGCCACCCCGTCATAAATTTACGATCTTTTATATGCAATCTCCAAAGCTGGGGCCAGGA 501
QY 142 GluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisValCys 161
DB 502 GAGAAGTATGAACTGCATGCGAGCGACAGACCACTCCAGTGTGGTGGTCCACGTGTGT 561
QY 162 GluSerAspGlnGluAsnGluGluGluGluGlu 173
DB 562 GAGAGTGACCAAGAGATGAGGAGGAGGAGGAGAGAG 597

RESULT 9

US-09-782-953-8
; Sequence 8, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)
; FILE REFERENCE: UTSD:674P21
; CURRENT APPLICATION NUMBER: US/09/782,953
; PRIOR FILING DATE: 2001-02-13
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 8
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(591)
US-09-782-953-8

Alignment Scores:

Pred. No.: 3.07e-66 Length: 594
Score: 613.50 Matches: 119

Percent Similarity:	75.00%	Conservative:	22
Best Local Similarity:	63.30%	Mismatches:	38
Query Match:	58.54%	Indels:	9
Ds:	10	Gaps:	2

; CURRENT APPLICATION NUMBER: US/09/782,953
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 3184
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (205)..(780)
US-09-782-953-17

Alignment Scores:
Pred. No.: 4,36e-63 Length: 3184
Score: 596.50 Matches: 117
Percent Similarity: 73.40% Conservatives: 21
Best Local Similarity: 62.23% Mismatches: 41
Query Match: 56.92% Indels: 9
DB: 10 Gaps: 2

US-09-782-953-3 (1-198) x US-09-782-953-17 (1-3184)

```
Qy 11 SerAlaThrIleAlaCysHisLeuAspProArgValPheValaspGlyLeuCysArgAla 30
Db 220 TCACACTGGTGGCTGCTGGTGGATGTCAGGCTCTTACCAATCAGGAGGTTAAGGAA 279
Qy 31 LysPheGluSerLeuPheArgThrTyrAspLysAspThrPheGlnTyrPheLysSer 50
Db 280 AATTTGGGGGACTGTTTCGACTATATGACTGTGTGACGTTCAGCTATTTAAGATT 339
Qy 51 PheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAspAlaArgLeuArg 70
Db 340 TTCAGAGCTGCTCGGTAAACTTCAGCAATCTTAAATCTGCAGCCGAGCTAGGATAGAG 399
Qy 71 LeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyrPheAlaGlnThrLeuHis 90
Db 400 CTTTCATGAACCCCAATTCAGAGGAAATAATTAAGCTCTACTTTGCACAGGTTTCAGACT 459
Qy 91 IleGlySerSer-----HisLeuAlaProProAsnProAspLysGlnPheLeu 106
Db 460 CCAGACAGACAGATGAGACAACTGCACTGGCTCCACCCAGGCTCCCAACAGTTTCTC 519
Qy 107 IleSerProProAlaSerProProValGlyTyrLysGlnValGluAspAlaThrProVal 126
Db 520 ATCTGCCCCCTTCTCTCCACCTGTTAGCTGGCAGCCCATCAACGATGCCAGCCAGTC 579
Qy 127 IleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyProGlyGluLysTyrGluLeu 146
Db 580 CTCNACTATGACTCTCTATGCTGTGGCCAACTAGGACCAAGAGAGATATGAGCTC 639
Qy 147 HisAlaAlaThrAspProThrProSerValValHisValCysGluSerAspGlnGlu 166
Db 640 CATGACGAGACTGAGTCCACCCCACTGCTGTCGACGTGTGCGACGTGCGACAGTACATAGAG 699
Qy 167 AsnGluGluGluGluGluMetGluArgMetLysArgProLysProLysIleGln 186
Db 700 GAAGAGAGAGGAGCAAAAG-----ACTTCCCAAGAGCAAAATCATCCAA 744
Qy 187 ThrArgAtgProGluTyrThrPro 194
Db 745 ACTCGGCTCTGGCTGCCACCC 768
```

RESULT 12

US-09-782-953-20
; Sequence 20, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE

; TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)
; FILE REFERENCE: UTSD:674P21
; CURRENT APPLICATION NUMBER: US/09/782,953
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 828
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (23)..(745)
US-09-782-953-20

Alignment Scores:
Pred. No.: 6,94e-64 Length: 828
Score: 596.00 Matches: 120
Percent Similarity: 72.16% Conservatives: 20
Best Local Similarity: 61.86% Mismatches: 44
Query Match: 56.87% Indels: 10
DB: 10 Gaps: 2

US-09-782-953-3 (1-198) x US-09-782-953-20 (1-828)

```
Qy 2 GluGluValAspLeuGlnAspLeuProSerAlaThrIleAlaCysHisLeuAspProArg 21
Db 134 GAGATGATGGATTTAAGTGATCTGCCTACCTCTTTTGTTCAGCGCTCCATGAAGCA 193
Qy 22 ValPheValAspGlyLeuCysArgAlaLysPheGluSerLeuPheArgThrTyrAspLys 41
Db 194 GTGTTTGAGGACGACGAGACAGAGAAAGATTGAAGCACTCTTCACCATCTATGATGAC 253
Qy 42 AspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsnPro 61
Db 254 CAGGTTACTTTTCAGCTGTTTAAAGCTTTAGAAGAGTCAGAAATAATTTTCAGCAAACT 313
Qy 62 LeuSerAlaAlaAspAlaArgLeuArgLysHisLysThrGluPheLeuGlyLysGluMet 81
Db 314 GAAGCGGACGACGAGCGCGAATAGAACTCCAGAAAACAGACTTCAATGGCGCAGAAGCTA 373
Qy 82 LysLeuTyrPheAlaGlnThrLeuHisIleGly-----SerSerHisLeuAla 97
Db 374 AAGCTATATTTTGCACAGGTCAGATGTCCGCGCAAGTGGCGGCAAGTCTATCTCTG 433
Qy 98 ProProAsnProAspLysGlnPheLeuLysSerProProAlaSerProProValGlyTyr 117
Db 434 CCGCCCCAGCCCTGTCAAGCAGTCTCTCATCTCCCTCCAGCTCTCCCCCAGTGGGTGG 493
Qy 118 LysGlnValGluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLys 137
Db 494 AAGCAGAGCGAAGATGCGATGCTGTTATAAATATGATTACTCTGTGTTTCCAAA 553
Qy 138 LeuGlyProGlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValVal 157
Db 554 TTGGGACCGAGGAGAAATATGAATTCACGCGGGAACAGAGTCGACACCCAGCGTGGTG 613
Qy 158 ValHisValCysGluSerAspGlnGluAsnGluGluGluGluMetGluArgMet 177
Db 614 GTTCATGCTGTGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 655
Qy 178 LysArgProLysProLysIleGlnThrArgArgProGlu 191
Db 656 AAAAAACCCCAAAACAGAAAAATTGCCACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 697
```

RESULT 13

US-09-782-953-23
; Sequence 23, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS

QY 36 AlaLysPheGluSerLeuPheArgThrTyrAspLysAspThrThrPheGlnTyrPheLys 49
Db 410 GCCAAATTGAGTCCCTCTTTAGGACGTATGACAAGGACATCACCTTTCAGTATTTTAAG 351
QY 50 SerPheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAapAlaArgLeu 69
Db 350 AGCTTCAACAGGTCAGATAAACTTCAGCAACCCCTTCTCCGAGCAGATGCCAGGCTC 291
QY 70 ArgLeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyrPheAlaGlnThrLeu 89
Db 290 CAGCTGCATAAGAGTGGTTCCTGGGAAAGGAATGAAGTATATTTGCTCAGGTGAGT 231
QY 90 HisIle-----GlySerSerHisLeuAlaProProAsn 100
Db 230 TGGGTTTCATTGCTATGATGCTTCTCTCCCTCTCCCTCCCTCCCTCC 189

RESULT 15

US-09-864-761-2064/c
; Sequence 2064, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Hanzel, David R.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeonica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 2064
; LENGTH: 446
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

; OTHER INFORMATION: MAP TO AP000122.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.1
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 4.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.4
US-09-864-761-2064

Alignment Scores:

Pred. No.:	2,79e-26	Length:	446
Score:	285.50	Matches:	58
Percent Similarity:	82.43%	Conservative:	3
Best Local Similarity:	78.38%	Mismatches:	10
Query Match:	27.24%	Indels:	3
DB:	10	Gaps:	1

US-09-782-953-3 (1-198) x US-09-864-761-2064 (1-446)

QY 30 AlaLysPheGluSerLeuPheArgThrTyrAspLysAspThrThrPheGlnTyrPheLys 49
Db 444 GCCAAATTGAGTCCCTCTTTAGGACGTATGACAAGGACATCACCTTTCAGTATTTTAAG 385
QY 50 SerPheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAapAlaArgLeu 69
Db 384 AGCTTCAACAGGTCAGATAAACTTCAGCAACCCCTTCTCCGAGCAGATGCCAGGCTC 325
QY 70 ArgLeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyrPheAlaGlnThrLeu 89
Db 324 CAGCTGCATAAGAGTGGTTCCTGGGAAAGGAATGAAGTATATTTGCTCAGGTGAGT 265
QY 90 HisIle-----GlySerSerHisLeuAlaProProAsn 100
Db 264 TGGGTTTCATTGCTATGATGCTTCTCTCCCTCTCCCTCCCTCCCTCC 223

Search completed: December 15, 2002, 00:06:20

Job time : 38.0556 secs

THIS PAGE BLANK (USPTO)


```

; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724.676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 61358
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-61358

```

Query Match	83.3%	Score	872.5	DB 5	Length	197			
Best Local Similarity	89.4%	Pred. No.	1.9e-81						
Matches	168	Conservative	7	Mismatches	12	Indels	1	Gaps	1
QY	11	SATIA	CHLDPRVFDGLCR	AFESLPTDYKDITTFQYFKSKVRINFSNPLSAADARLR	70				
		:	:	:	:	:	:	:	:
Db	11	SSLIACVANSIDFSE	TRAFESLFRUYRDKDITTFQYFKSKVRINFSNPFSAADARLQ	70					
QY	71	LHKT	BFLGKEMKLYPAQ	TLHGSSHLAPPNDPKOFLISPPASPVGWKQVEDATPVIND	130				
		:	:	:	:	:	:	:	:
Db	71	LHKT	BFLGKEMKLYPAQ	TLHGSSHLAPPNDPKOFLISPPASPVGWKQVEDATPVIND	130				
QY	131	LYIA	LSKLGCEKYE	LHAATPTSPVVHVCSDOENEEBEEEMERMKPKKIQTTRP	190				
		:	:	:	:	:	:	:	:
Db	131	LYIA	LSKLGCEKYE	LHAATPTSPVVHVCSDOE-KEEEEMERMRPKKIQTTRP	189				
QY	191	EYTP	IHL	S	198				
		:	:	:	:	:	:	:	:
Db	190	EYTP	IHL	S	197				

RESULT 7
US-09-724-676A-61359
; Sequence 61359, Application US/09724676A

```
Query Match      83.3%; Score 872.5; DB 5; Length 197;
Beat Local Similarity 89.4%; Pred. No. 1.9e-81;
Matches 168; Conservative 7; Mismatches 12; Indels 1; Gaps 1;
```

OY	11 SATIACHLDPRVFVDGLCRKFESLFETYDKDITFOYKSFKRVRINFNSNPUSAADARLQ	70
Ddb	: : : : : : : : : : : : : : : : : :	:
OY	11 SSLIACVANSIDIFSESTRAKFESLFETYDKDITFOYKSFKRVRINFNSNPFSAADARLQ	70
Ddb	:	:
OY	71 LHKTEFLGCKMKLYFAOTLHGSHLLAPPNDPKQFLISPPASPPVGWKQVEDATPVIND	130
Ddb	:	:
OY	71 LHKTEFLGCKMKLYFAOTLHGSHLLAPPNDPKQFLISPPASPPVGWKQVEDATPVIND	130
OY	131 LLAYATSKLGPCKEYELHAAATPTSPVVHVHCSDQNEEEEEEMERMRKPPIQTTRP	190
Ddb	: : : : : : : : : : : : : :	:
OY	131 LLAYATSKLGPCKEYELHAAATPTSPVVHVHCSDQE-KSEEBEEMERMRRPKKIQTTRP	189
OY	191 EYTPIHLS	198
Ddb	: : : : : : : :	:
OY	190 EYTPIHLS	197

RESULT 8
US-09-724-676A-61367

```

: Sequence 61367, Application US/09724676A
:
: GENERAL INFORMATION:
: APPLICANT: Comugen LTD
: TITLE OF INVENTION: Variants of alternative splicing
: FILE REFERENCE: 123181.4 Comugen
: CURRENT APPLICATION NUMBER: US/09/724, 676A
: CURRENT FILING DATE: 2000-11-28
: NUMBER OF SEQ ID NOS: 97222
: SOFTWARE: PatentIn version 3.2
: SEQ ID NO 61367
: LENGTH: 197
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-724-676A-61367

```

Query Match	83.3%	Score	872.5	DB 5	Length	197
Best Local Similarity	89.4%	Pred. No.	1.9e-81			
Matches	168	Conservative	7	Mismatches	12	Indels
						Gaps
Qy	11	SATTACHLDPRVVDGLCRAKFESLFRTYDKDITTFQYFKSPKRVIRNFSNPLSAAADARLR	70			
Db	11	SSLIACVANGSDISESTRAKFESLFRTYDKDITTFQYFKSPKRVIRNFSNPFSAADARLQ	70			
Qy	71	LHKTFFLGKMKLYFAQTLHGSSHLAPPNDPKOFLISPPASPPVGMKQVEDATPVIYND	130			
Db	71	LHKTFFLGKMKLYFAQTLHGSSHLAPPNDPKOFLISPPASPPVGMKQVEDATPVIYND	130			
Qy	131	LLVAISKLGPEGYELHAATDPTSPVVHVHCESDOENEESEEMERMKRPKPKIIQTRRP	190			
Db	131	LLVAISKLGPEGYELHAATDPTSPVVHVHCESDOE-KEEESEEMRRRPKPKIIQTRRP	189			
Qy	191	EYTPIHLS	198			
Db	190	EYTPIHLS	197			

RESULT 9
US-09-724-676-61356
; Sequence 61356, Application US/09724676

Query Match		80.9%;	Score 848;	DB 5;	Length 200;	
Best Local Similarity		85.6%;	Pred. No. 6e-79;			
Matches 167; Conservative		9;	Mismatches 13;	Indels 6;	Gaps 3;	
QY	4	VLDQLPSATIAICHLDPFVDCGLCRAKTESLFTYDKDTTFQYFKSPKVRINFNSPLS	63			
		: :				
Dd	12	INEDDVAPFI -CTIGKANV-----WAKESLFTYDKDITTFQYFKSPKVRINFNSPFS	66			
QY	64	ADARURLHKTFLGKEMKLYFAOTLIHGSHLAPPNDPKOFLISPPASPVGWKVQEDA	123			
		: :				
Dd	67	AADARLQLHKTEFLGKEMKLYFAOTLIHGSHLAPPNDPKOFLISPPASPVGWKVQEDA	126			
QY	124	TPVINVDLLYAISKLGPGKEYELHAATDPTPSVVHVCHSDSENEEEEMERMKRKKPK	183			
		: :				
Dd	127	TPVINVDLLYAISKLGPGKEYELHAATDTPSVVHVCHSDQB-KGESEEMRRMRKKPK	185			
QY	184	IIOTRREPYTPIHLS	198			
Dd	186	IIOTRREPYTPIHLS	200			

;; TITLE OF INVENTION: Variants of alternative splicing
;; FILE REFERENCE: 129181.4 Compugen
;; CURRENT APPLICATION NUMBER: US/09/724,676A
;; CURRENT FILING DATE: 2000-11-28
;; NUMBER OF SEQ ID NOS: 97222
;; SOFTWARE: PatentIn version 3.2
;; SEQ ID NO 61353
;; LENGTH: 154
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-724-676A-61353

Query Match 69.8%; Score 731; DB 5; Length 154;
Best Local Similarity 95.2%; Pred. No. 3.7e-67;
Matches 138; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 1 MEEVDLQDLPSATIACHLDPFVVDGLCRAKFESLFRDYDKDTTFQYFKSFKRVIRNFSN 60
DB 1 MEEVDLEDLPSATIACHLDPFVVDGLCRAKFESLFRDYDKDTTFQYFKSFKRVIRNFSN 60
QY 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPPVGWKQV 120
DB 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPPVGWKQV 120
QY 121 EDATPVINYDLLYAIKLGGEKYE 145
DB 121 EDATPVINYDLLYAIKLGGEKYE 145

RESULT 15
US-10-290-438-2
;; Sequence 2, Application US/10290438
;; GENERAL INFORMATION:
;; APPLICANT: Loring, Jeanne F.
;; APPLICANT: Tingley, Debora W.
;; APPLICANT: Edwards, Carla M.
;; APPLICANT: Streeter, David G.
;; TITLE OF INVENTION: DOWN SYNDROME CRITICAL REGION 1-LIKE PROTEIN
;; FILE REFERENCE: FC-0013-1C1P
;; CURRENT APPLICATION NUMBER: US/10/290,438
;; CURRENT FILING DATE: 2002-11-06
;; PRIOR APPLICATION NUMBER: 09/614,474
;; PRIOR FILING DATE: 2000-07-11
;; NUMBER OF SEQ ID NOS: 11
;; SOFTWARE: PERL Program
;; SEQ ID NO 2
;; LENGTH: 255
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc_feature
;; OTHER INFORMATION: Incyte ID No: 247500.5
;; FEATURE:
;; NAME/KEY: unsure
;; LOCATION: 11
;; OTHER INFORMATION: unknown or other
US-10-290-438-2

Query Match 59.2%; Score 620.5; DB 6; Length 255;
Best Local Similarity 62.4%; Pred. No. 1.3e-55;
Matches 121; Conservative 22; Mismatches 42; Indels 9; Gaps 2;
QY 5 DLQDLPSATIACHLDPFVVDGLCRAKFESLFRDYDKDTTFQYFKSFKRVIRNFSNPLSA 64
DB 63 DFNDLPNSLFACNVHQSVEGESEKSEKGLFRDYDDCVTFQLFKSFRRVIRNFSNPKSA 122
QY 65 ADARLRLHKTFLGKEMKLYFAQTLHGSS----HLAPPNDKQFLISPPASPPVGWKQV 120
DB 123 ADARLRLHKTFLGKEMKLYFAQTLHGSS----HLAPPNDKQFLISPPASPPVGWKQV 120
QY 121 EDATPVINYDLLYAIKLGGEKYEELHAATDTPSVVHVCSDOENEEEMERMKRP 180
DB 183 NDATPVINYDLLYAVAKLGGEKYEELHAGTESTPSVVHVCDSDIEEDDPK-----TSP 237

QY 181 KPKIIQTRRPEYTP 194
DB 238 KPKIIQTRRPGLP 251
Search completed: December 11, 2002, 11:50:14
Job time : 12.6863 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 11:39:49 ; Search time 6.34401 Seconds
(without alignments)
506.931 Million cell updates/sec

Title: US-09-782-953-3
Perfect score: 1048
Sequence: 1 MEEVDLQDLPSTAIACHLDP.....RPKIIQTRPEYTPIHLS 198

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 103943 seqs, 16242309 residues

Total number of hits satisfying chosen parameters: 103943

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pdb.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pdb.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pdb.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pdb.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pdb.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pdb.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pdb.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pdb.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pdb.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pdb.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pdb.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pdb.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pdb.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pdb.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1048	100.0	198	10	US-09-782-953-3
2	1048	100.0	198	10	US-09-782-953-4
3	1002.5	95.7	197	10	US-09-782-953-15
4	1002.5	95.7	197	10	US-09-782-953-16
5	920.5	87.8	198	10	US-09-782-953-6
6	920.5	87.8	198	10	US-09-782-953-7
7	872.5	83.3	197	10	US-09-782-953-12
8	872.5	83.3	197	10	US-09-782-953-13
9	613.5	58.5	197	10	US-09-782-953-9
10	613.5	58.5	197	10	US-09-782-953-10
11	596.5	56.9	192	10	US-09-782-953-18
12	596.5	56.9	192	10	US-09-782-953-19
13	596.5	56.9	241	10	US-09-782-953-21
14	596.5	56.9	241	10	US-09-782-953-22
15	595.5	56.8	212	10	US-09-782-953-24
16	595.5	56.8	212	10	US-09-782-953-25
17	497.4	47.4	142	10	US-09-325-302-790
18	284	27.1	58	10	US-09-864-761-35379
19	284	27.1	58	10	US-09-864-761-43076

20	271.5	25.9	56	10	US-09-864-761-34111	Sequence 34111, A
21	90.5	8.6	4019	10	US-09-738-973-425	Sequence 425, App
22	86.5	8.3	822	10	US-09-824-734-3	Sequence 3, Appli
23	79	7.5	572	10	US-09-771-161A-120	Sequence 120, App
24	79	7.5	1142	10	US-09-771-161A-211	Sequence 211, App
25	78.5	7.5	864	10	US-09-883-096-2	Sequence 2, Appli
26	77	7.3	362	9	US-09-349-755-32	Sequence 32, Appli
27	77	7.3	362	9	US-09-166-334-32	Sequence 32, Appli
28	77	7.3	362	10	US-09-350-206-32	Sequence 32, Appli
29	76	7.3	445	9	US-09-349-755-5	Sequence 5, Appli
30	76	7.3	445	9	US-09-166-334-5	Sequence 5, Appli
31	76	7.3	445	10	US-09-350-206-5	Sequence 5, Appli
32	75	7.2	195	9	US-09-902-941-1888	Sequence 1888, Ap
33	75	7.2	1885	10	US-09-920-346-2	Sequence 2, Appli
34	74	7.1	311	10	US-09-819-252-2	Sequence 2, Appli
35	74	7.1	311	10	US-09-922-217-1061	Sequence 1061, Ap
36	74	7.1	311	10	US-09-833-263-1061	Sequence 1061, Ap
37	73.5	7.0	536	10	US-09-908-805B-61	Sequence 61, Appli
38	72.5	6.9	536	10	US-09-908-805B-18	Sequence 18, Appli
39	72.5	6.9	559	10	US-09-854-549-7	Sequence 7, Appli
40	72.5	6.9	559	10	US-09-836-561-6	Sequence 6, Appli
41	72.5	6.9	1924	9	US-09-866-557A-2	Sequence 2, Appli
42	72	6.9	591	10	US-09-887-586A-24	Sequence 24, Appli
43	72	6.9	591	10	US-09-903-012-24	Sequence 24, Appli
44	71.5	6.8	284	10	US-09-764-864-843	Sequence 843, App
45	71.5	6.8	590	10	US-09-815-242-10812	Sequence 10812, A

ALIGNMENTS

RESULT 1
US-09-782-953-3
; Sequence 3, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)
; CURRENT APPLICATION NUMBER: US/09/782,953
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 3
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-782-953-3

Query Match 100.0%; Score 1048; DB 10; Length 198;
Best Local Similarity 100.0%; Pred. No. 3.5e-99;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MEEVDLQDLPSTAIACHLDPVFVDGLCRKPFSLPRTYDKDTTFQYFKSFKRVINFSN 60
Db 1 MEEVDLQDLPSTAIACHLDPVFVDGLCRKPFSLPRTYDKDTTFQYFKSFKRVINFSN 60
Qy 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHGSSHLAPPNDPKQFLISPPASPPVGWQV 120
Db 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHGSSHLAPPNDPKQFLISPPASPPVGWQV 120
Qy 121 EDATPVINYDLLLVAISKLGEGKVELHAATDPTPSVVHVHVCSDQNEEEEEEMERKRP 180
Db 121 EDATPVINYDLLLVAISKLGEGKVELHAATDPTPSVVHVHVCSDQNEEEEEEMERKRP 180
Qy 181 KPKIIQTRPEYTPIHLS 198
Db 181 KPKIIQTRPEYTPIHLS 198

```
RESULT 2
US-09-782-953-4
; Sequence 4, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674P21
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: US/09/782,953
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-782-953-4

Query Match
Best Local Similarity 100.0%; Score 1048; DB 10; Length 198;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEEVDLQDLPSATIACHLDPRVFDGLCRAKFESLFRITYDKDTTFQYFKSPKRVIRINFSN 60
DB 1 MEEVDLQDLPSATIACHLDPRVFDGLCRAKFESLFRITYDKDTTFQYFKSPKRVIRINFSN 60
QY 61 PLSAADARLRLHKTETFLGKEMKLYFAQTLHIGSSHLAPPNDKQFLISPPASPPVGMKQV 120
DB 61 PLSAADARLRLHKTETFLGKEMKLYFAQTLHIGSSHLAPPNDKQFLISPPASPPVGMKQV 120
QY 121 EDATPVINYDLVAISKLGPEKYEELHAATDPTPSVVHVHVCSDQENEEBEMERMRKP 180
DB 121 EDATPVINYDLVAISKLGPEKYEELHAATDPTPSVVHVHVCSDQENEEBEMERMRKP 180
QY 181 KPKIITQTRRPETPIHLS 198
DB 181 KPKIITQTRRPETPIHLS 198

RESULT 3
US-09-782-953-15
; Sequence 15, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674P21
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: US/09/782,953
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 15
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-15

Query Match
Best Local Similarity 95.7%; Score 1002.5; DB 10; Length 197;
Matches 191; Conservative 3; Mismatches 3; Indels 1; Gaps 1;
```

```
QY 61 PLSAADARLRLHKTETFLGKEMKLYFAQTLHIGSSHLAPPNDKQFLISPPASPPVGMKQV 120
DB 61 PLSAADARLRLHKTETFLGKEMKLYFAQTLHIGSSHLAPPNDKQFLISPPASPPVGMKQV 120
QY 121 EDATPVINYDLVAISKLGPEKYEELHAATDPTPSVVHVHVCSDQENEEBEMERMRKP 180
DB 121 EDATPVINYDLVAISKLGPEKYEELHAATDPTPSVVHVHVCSDQENEEBEMERMRKP 180
QY 181 KPKIITQTRRPETPIHLS 198
DB 181 KPKIITQTRRPETPIHLS 197

RESULT 4
US-09-782-953-16
; Sequence 16, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674P21
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: US/09/782,953
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 16
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-16

Query Match
Best Local Similarity 95.7%; Score 1002.5; DB 10; Length 197;
Matches 191; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

QY 1 MEEVDLQDLPSATIACHLDPRVFDGLCRAKFESLFRITYDKDTTFQYFKSPKRVIRINFSN 60
DB 1 MEEVDLQDLPSATIACHLDPRVFDGLCRAKFESLFRITYDKDTTFQYFKSPKRVIRINFSN 60
QY 61 PLSAADARLRLHKTETFLGKEMKLYFAQTLHIGSSHLAPPNDKQFLISPPASPPVGMKQV 120
DB 61 PLSAADARLRLHKTETFLGKEMKLYFAQTLHIGSSHLAPPNDKQFLISPPASPPVGMKQV 120
QY 121 EDATPVINYDLVAISKLGPEKYEELHAATDPTPSVVHVHVCSDQENEEBEMERMRKP 180
DB 121 EDATPVINYDLVAISKLGPEKYEELHAATDPTPSVVHVHVCSDQENEEBEMERMRKP 180
QY 181 KPKIITQTRRPETPIHLS 198
DB 181 KPKIITQTRRPETPIHLS 197

RESULT 5
US-09-782-953-6
; Sequence 6, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674P21
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: US/09/782,953
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin Ver. 2.1
```

```
; SEQ ID NO 6
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-782-953-6

Query Match      87.8%; Score 920.5; DB 10; Length 198;
Best Local Similarity 89.9%; Pred. No. 3.1e-86;
Matches 178; Conservative 5; Mismatches 12; Indels 3; Gaps 1;

QY 4 VLDQLP---SATIACHLDPRVVDGLCRAKFESLRTYDKDTTFQYFKSFKRVINFSN 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 MDRDFSYNFSLLIACVANDVSESETRAKFESLRTYDKDTTFQYFKSFKRVINFSN 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHIGSSHLAPPNDPKQFLISPPASPPVGKQV 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHIGSSHLAPPNDPKQFLISPPASPPVGKQV 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 121 EDATPVINYDLLYAIKLGPGKEKLYELHAATDTPSVVVHVCSDDQNEEEEEEMERMKRP 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 121 EDATPVINYDLLYAIKLGPGKEKLYELHAATDTPSVVVHVCSDDQNEEEEEEMERMKRP 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 181 KPKIOTRRPEYTPHLS 198
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 181 KPKIOTRRPEYTPHLS 198
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 6
US-09-782-953-7
; Sequence 7, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674P21
; CURRENT APPLICATION NUMBER: US/09/782,953
; PRIOR FILING DATE: 2001-02-13
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-782-953-7

Query Match      87.8%; Score 920.5; DB 10; Length 198;
Best Local Similarity 89.9%; Pred. No. 3.1e-86;
Matches 178; Conservative 5; Mismatches 12; Indels 3; Gaps 1;

QY 4 VLDQLP---SATIACHLDPRVVDGLCRAKFESLRTYDKDTTFQYFKSFKRVINFSN 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 MDRDFSYNFSLLIACVANDVSESETRAKFESLRTYDKDTTFQYFKSFKRVINFSN 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHIGSSHLAPPNDPKQFLISPPASPPVGKQV 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHIGSSHLAPPNDPKQFLISPPASPPVGKQV 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 121 EDATPVINYDLLYAIKLGPGKEKLYELHAATDTPSVVVHVCSDDQNEEEEEEMERMKRP 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 121 EDATPVINYDLLYAIKLGPGKEKLYELHAATDTPSVVVHVCSDDQNEEEEEEMERMKRP 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 181 KPKIOTRRPEYTPHLS 198
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 181 KPKIOTRRPEYTPHLS 198
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 7
US-09-782-953-12
; Sequence 12, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674P21
; CURRENT APPLICATION NUMBER: US/09/782,953
; PRIOR FILING DATE: 2001-02-13
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-12

Query Match      83.3%; Score 872.5; DB 10; Length 197;
Best Local Similarity 89.4%; Pred. No. 2.3e-81;
Matches 168; Conservative 7; Mismatches 12; Indels 1; Gaps 1;

QY 11 SATIACHLDPRVVDGLCRAKFESLRTYDKDTTFQYFKSFKRVINFSNPLSADARLQ 70
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 11 SSLIACVANSIDFSESETRAKFESLRTYDKDTTFQYFKSFKRVINFSNPLSADARLQ 70
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 71 LHKTEFLGKEMKLYFAQTLHIGSSHLAPPNDPKQFLISPPASPPVGKQVEDATPVINYD 130
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 71 LHKTEFLGKEMKLYFAQTLHIGSSHLAPPNDPKQFLISPPASPPVGKQVEDATPVINYD 130
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 131 LLYAISKLGPGKEKLYELHAATDTPSVVVHVCSDDQNEEEEEEMERMKPKKIOTRRP 190
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 131 LLYAISKLGPGKEKLYELHAATDTPSVVVHVCSDDQNEEEEEEMERMKPKKIOTRRP 190
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 191 EYTPHLS 198
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 190 EYTPHLS 197
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 8
US-09-782-953-13
; Sequence 13, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674P21
; CURRENT APPLICATION NUMBER: US/09/782,953
; CURRENT FILING DATE: 2001-02-13
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-13

Query Match      83.3%; Score 872.5; DB 10; Length 197;
Best Local Similarity 89.4%; Pred. No. 2.3e-81;
Matches 168; Conservative 7; Mismatches 12; Indels 1; Gaps 1;

QY 11 SATIACHLDPRVVDGLCRAKFESLRTYDKDTTFQYFKSFKRVINFSNPLSADARLQ 70
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 11 SSLIACVANSIDFSESETRAKFESLRTYDKDTTFQYFKSFKRVINFSNPLSADARLQ 70
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 71 LHKTEFLGKEMKLYFAQTLHIGSSHLAPPNDPKQFLISPPASPPVGKQVEDATPVINYD 130
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 71 LHKTEFLGKEMKLYFAQTLHIGSSHLAPPNDPKQFLISPPASPPVGKQVEDATPVINYD 130
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

; SEQ ID NO 6
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-782-953-6

Query Match      87.8%; Score 920.5; DB 10; Length 198;
Best Local Similarity 89.9%; Pred. No. 3.1e-86;
Matches 178; Conservative 5; Mismatches 12; Indels 3; Gaps 1;

QY 4 VLDQLP---SATIACHLDPRVVDGLCRAKFESLRTYDKDTTFQYFKSFKRVINFSN 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 MDRDFSYNFSLLIACVANDVSESETRAKFESLRTYDKDTTFQYFKSFKRVINFSN 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHIGSSHLAPPNDPKQFLISPPASPPVGKQV 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHIGSSHLAPPNDPKQFLISPPASPPVGKQV 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 121 EDATPVINYDLLYAIKLGPGKEKLYELHAATDTPSVVVHVCSDDQNEEEEEEMERMKRP 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 121 EDATPVINYDLLYAIKLGPGKEKLYELHAATDTPSVVVHVCSDDQNEEEEEEMERMKRP 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 181 KPKIOTRRPEYTPHLS 198
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 181 KPKIOTRRPEYTPHLS 198
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 6
US-09-782-953-7
; Sequence 7, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674P21
; CURRENT APPLICATION NUMBER: US/09/782,953
; PRIOR FILING DATE: 2001-02-13
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-782-953-7

Query Match      87.8%; Score 920.5; DB 10; Length 198;
Best Local Similarity 89.9%; Pred. No. 3.1e-86;
Matches 178; Conservative 5; Mismatches 12; Indels 3; Gaps 1;

QY 4 VLDQLP---SATIACHLDPRVVDGLCRAKFESLRTYDKDTTFQYFKSFKRVINFSN 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 MDRDFSYNFSLLIACVANDVSESETRAKFESLRTYDKDTTFQYFKSFKRVINFSN 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHIGSSHLAPPNDPKQFLISPPASPPVGKQV 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHIGSSHLAPPNDPKQFLISPPASPPVGKQV 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 121 EDATPVINYDLLYAIKLGPGKEKLYELHAATDTPSVVVHVCSDDQNEEEEEEMERMKRP 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 121 EDATPVINYDLLYAIKLGPGKEKLYELHAATDTPSVVVHVCSDDQNEEEEEEMERMKRP 180
   : : : : ~~~~~~
```

Qy 131 LLYAISKLGPEGEKYLHAATDPTPSVVHVHVCSDQNEEEEMERMRKPKKXIOTRRP 190
 Db 131 LLYAISKLGPEGEKYLHAATDPTPSVVHVHVCSDQNEEEEMERMRKPKKXIOTRRP 189
 Qy 191 EYTPHLS 198
 Db 190 EYTPHLS 197

RESULT 9
 US-09-782-953-9
 ; Sequence 9, Application US/09782953
 ; Patent No. US20020150953A1
 ; GENERAL INFORMATION:
 ; APPLICANT: WILLIAMS, R. SANDERS
 ; APPLICANT: ROTHERMEL, BEVERLY
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
 ; TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)
 ; FILE REFERENCE: UTSD:674P21
 ; CURRENT APPLICATION NUMBER: US/09/782,953
 ; CURRENT FILING DATE: 2001-02-13
 ; PRIOR APPLICATION NUMBER: 60/216,601
 ; PRIOR FILING DATE: 2000-07-07
 ; NUMBER OF SEQ ID NOS: 27
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 9
 ; LENGTH: 197
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 US-09-782-953-9

Query Match 58.5%; Score 613.5; DB 10; Length 197;
 Best Local Similarity 63.3%; Pred. No. 4.5e-55;
 Matches 119; Conservative 22; Mismatches 38; Indels 9; Gaps 2;

Qy 11 SATIACHLDPRVVDGLCRAKFESLFRITDKDTTFQYFKSFKRVIRINFSNPLSADARLR 70
 Db 11 STLVACVVDVEVFTNOEVEKEFGFLFRITDECVTFLFKSFRRVIRINFSNPKSAARARIE 70
 Qy 71 LKTEFLGKEMKLYFAQTLHIGSS----HLAPPNDKQFLISPPASPPVGMKQVEDATPV 126
 Db 71 LHETQGRGKKLKYFAQVOTPETDGDKHLAPQPAKQFLISPPSSPPVGMKPISDATPV 130
 Qy 127 INVYLLYAISKLGPEGEKYLHAATDPTPSVVHVHVCSDQNEEEEMERMRKPKKXIIO 186
 Db 131 LNYDLYAVAKLGPEGEKYLHAGTSTPSVVHVHVCSDMEEBEDPK-----TSPKPKXIIO 185

Qy 187 TRPEYTP 194
 Db 186 TRPGLPP 193

RESULT 10
 US-09-782-953-10
 ; Sequence 10, Application US/09782953
 ; Patent No. US20020150953A1
 ; GENERAL INFORMATION:
 ; APPLICANT: WILLIAMS, R. SANDERS
 ; APPLICANT: ROTHERMEL, BEVERLY
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
 ; TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)
 ; FILE REFERENCE: UTSD:674P21
 ; CURRENT APPLICATION NUMBER: US/09/782,953
 ; CURRENT FILING DATE: 2001-02-13
 ; PRIOR APPLICATION NUMBER: 60/216,601
 ; PRIOR FILING DATE: 2000-07-07
 ; NUMBER OF SEQ ID NOS: 27
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 10
 ; LENGTH: 197
 ; TYPE: PRT
 ; ORGANISM: Mus musculus

US-09-782-953-10

Query Match 58.5%; Score 613.5; DB 10; Length 197;
 Best Local Similarity 63.3%; Pred. No. 4.5e-55;
 Matches 119; Conservative 22; Mismatches 38; Indels 9; Gaps 2;

Qy 11 SATIACHLDPRVVDGLCRAKFESLFRITDKDTTFQYFKSFKRVIRINFSNPLSADARLR 70
 Db 11 STLVACVVDVEVFTNOEVEKEFGFLFRITDECVTFLFKSFRRVIRINFSNPKSAARARIE 70
 Qy 71 LKTEFLGKEMKLYFAQTLHIGSS----HLAPPNDKQFLISPPASPPVGMKQVEDATPV 126
 Db 71 LHETQGRGKKLKYFAQVOTPETDGDKHLAPQPAKQFLISPPSSPPVGMKPISDATPV 130
 Qy 127 INVYLLYAISKLGPEGEKYLHAATDPTPSVVHVHVCSDQNEEEEMERMRKPKKXIIO 186
 Db 131 LNYDLYAVAKLGPEGEKYLHAGTSTPSVVHVHVCSDMEEBEDPK-----TSPKPKXIIO 185
 Qy 187 TRPEYTP 194
 Db 186 TRPGLPP 193

RESULT 11
 US-09-782-953-18
 ; Sequence 18, Application US/09782953
 ; Patent No. US20020150953A1
 ; GENERAL INFORMATION:
 ; APPLICANT: WILLIAMS, R. SANDERS
 ; APPLICANT: ROTHERMEL, BEVERLY
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
 ; TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)
 ; FILE REFERENCE: UTSD:674P21
 ; CURRENT APPLICATION NUMBER: US/09/782,953
 ; CURRENT FILING DATE: 2001-02-13
 ; PRIOR APPLICATION NUMBER: 60/216,601
 ; PRIOR FILING DATE: 2000-07-07
 ; NUMBER OF SEQ ID NOS: 27
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 18
 ; LENGTH: 192
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-782-953-18

Query Match 56.9%; Score 596.5; DB 10; Length 192;
 Best Local Similarity 62.2%; Pred. No. 2.3e-53;
 Matches 117; Conservative 21; Mismatches 41; Indels 9; Gaps 2;

Qy 11 SATIACHLDPRVVDGLCRAKFESLFRITDKDTTFQYFKSFKRVIRINFSNPLSADARLR 70
 Db 6 STLVACVVDVEVFTNOEVEKEFGFLFRITDDCVTFQFLFKSFRRVIRINFSNPKSAARARIE 65
 Qy 71 LKTEFLGKEMKLYFAQTLHIGSS----HLAPPNDKQFLISPPASPPVGMKQVEDATPV 126
 Db 66 LHETQGRGKKLKYFAQVOTPETDGDKHLAPQPAKQFLISPPSSPPVGMKPISDATPV 125
 Qy 127 INVYLLYAISKLGPEGEKYLHAATDPTPSVVHVHVCSDQNEEEEMERMRKPKKXIIO 186
 Db 126 LNYDLYAVAKLGPEGEKYLHAGTSTPSVVHVHVCSDMEEBEDPK-----TSPKPKXIIO 180
 Qy 187 TRPEYTP 194
 Db 181 TRPGLPP 188

RESULT 12
 US-09-782-953-19
 ; Sequence 19, Application US/09782953
 ; Patent No. US20020150953A1
 ; GENERAL INFORMATION:
 ; APPLICANT: WILLIAMS, R. SANDERS
 ; APPLICANT: ROTHERMEL, BEVERLY


```
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674PZ1
; CURRENT APPLICATION NUMBER: US/09/782,953
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-19

Query Match          56.9%; Score 596.5; DB 10; Length 192;
Best Local Similarity 62.2%; Pred. No. 2.3e-53;
Matches 117; Conservative 21; Mismatches 41; Indels 9; Gaps 2;

QY 11 SATIACHLDRPVFDGLCRAKFSLRTYDKDTTFQYFKSFKVRINFSNPLSAADARLR 70
DB 11 :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
6 STLVACVVDVEFTNQBKFGGLFRTYDDCVTFQLFKSFRRVRINFSNPKSAARARIE 65
QY 71 LHKTEFLGKMKLYFAQTLHGSS----HLAPPNDKQFLISPPASPPVCMKQVEDATPV 126
DB 71 :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
66 LHETQFGKKLYFAQVQTPETDQKHLHAPPQAKQFLISPPSPVSWQPINATPV 125
QY 127 INVYLLVAISKLGPGEKYLHAATDTPSVVHVHVCSDQNEEEEEEMERMKRKPQKIIQ 186
DB 127 :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
126 LNYDLLYAVAKLGPGEKYLHAGTSTPSPVVHVHVCSDIEEDPK-----TSPKPKIIQ 180
QY 187 TRPEYTP 194
DB 187 :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
181 TRPGUUP 188

RESULT 13
US-09-782-953-21
; Sequence 21, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674PZ1
; CURRENT APPLICATION NUMBER: US/09/782,953
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-21

Query Match          56.9%; Score 596; DB 10; Length 241;
Best Local Similarity 61.9%; Pred. No. 3.6e-53;
Matches 120; Conservative 20; Mismatches 44; Indels 10; Gaps 2;

QY 2 BEVDLQDLPSATIACHLDRPVFDGLCRAKFSLRTYDKDTTFQYFKSFKVRINFSNP 61
DB 2 :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
38 EMDLSDLPTSLFACSVHAEVFEAREQKRFALFTIYDDQVTFQLFKSFRRVRINFSKP 97
QY 62 LSAADARLRLHKTFLGKMKLYFAQTLHG----SSHLAPPNDKQFLISPPASPPVCGW 117
DB 62 :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
98 EAAARARIELHETDQKGLYFAQVQMSGEVRDKSYLLPPQPVKQFLISPPASPPVCGW 157
QY 118 KQVEDATPVINYDLLYVAISKLGPGEKYLHAATDTPSVVHVHVCSDQNEEEEEEMERM 177
DB 118 :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
158 KQSDAMPVINYDLLCAVSKLGPGEKYLHAGTSTPSPVVHVHVCSETEEBE-----T 211

US-09-782-953-22
; Sequence 22, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674PZ1
; CURRENT APPLICATION NUMBER: US/09/782,953
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-22

Query Match          56.9%; Score 596; DB 10; Length 241;
Best Local Similarity 61.9%; Pred. No. 3.6e-53;
Matches 120; Conservative 20; Mismatches 44; Indels 10; Gaps 2;

QY 2 BEVDLQDLPSATIACHLDRPVFDGLCRAKFSLRTYDKDTTFQYFKSFKVRINFSNP 61
DB 2 :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
38 EMDLSDLPTSLFACSVHAEVFEAREQKRFALFTIYDDQVTFQLFKSFRRVRINFSKP 97
QY 62 LSAADARLRLHKTFLGKMKLYFAQTLHG----SSHLAPPNDKQFLISPPASPPVCGW 117
DB 62 :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
98 EAAARARIELHETDQKGLYFAQVQMSGEVRDKSYLLPPQPVKQFLISPPASPPVCGW 157
QY 118 KQVEDATPVINYDLLYVAISKLGPGEKYLHAATDTPSVVHVHVCSDQNEEEEEEMERM 177
DB 118 :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
158 KQSDAMPVINYDLLCAVSKLGPGEKYLHAGTSTPSPVVHVHVCSETEEBE-----T 211

US-09-782-953-23
; Sequence 23, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674PZ1
; CURRENT APPLICATION NUMBER: US/09/782,953
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-23

Query Match          56.8%; Score 595; DB 10; Length 212;
Best Local Similarity 62.6%; Pred. No. 3.8e-53;
```

```
QY 178 KRPKKIIQTRRPE 191
DB 212 KNPKQKIAQTRRPD 225

RESULT 14
US-09-782-953-22
; Sequence 22, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674PZ1
; CURRENT APPLICATION NUMBER: US/09/782,953
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-22

Query Match          56.9%; Score 596; DB 10; Length 241;
Best Local Similarity 61.9%; Pred. No. 3.6e-53;
Matches 120; Conservative 20; Mismatches 44; Indels 10; Gaps 2;

QY 2 BEVDLQDLPSATIACHLDRPVFDGLCRAKFSLRTYDKDTTFQYFKSFKVRINFSNP 61
DB 2 :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
38 EMDLSDLPTSLFACSVHAEVFEAREQKRFALFTIYDDQVTFQLFKSFRRVRINFSKP 97
QY 62 LSAADARLRLHKTFLGKMKLYFAQTLHG----SSHLAPPNDKQFLISPPASPPVCGW 117
DB 62 :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
98 EAAARARIELHETDQKGLYFAQVQMSGEVRDKSYLLPPQPVKQFLISPPASPPVCGW 157
QY 118 KQVEDATPVINYDLLYVAISKLGPGEKYLHAATDTPSVVHVHVCSDQNEEEEEEMERM 177
DB 118 :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
158 KQSDAMPVINYDLLCAVSKLGPGEKYLHAGTSTPSPVVHVHVCSETEEBE-----T 211

US-09-782-953-24
; Sequence 24, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674PZ1
; CURRENT APPLICATION NUMBER: US/09/782,953
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 24
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-24

Query Match          56.8%; Score 595; DB 10; Length 212;
Best Local Similarity 62.6%; Pred. No. 3.8e-53;
```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 11:38:44 ; Search time 136.563 Seconds
(without alignments)
934.784 Million cell updates/sec

Title: US-09-782-953-3
Perfect score: 1048
Sequence: 1 MEEVDQLPSATIACHLDP.....RPKPIQTRPEYTPHLS 198

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents AA Main:*

- 1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep.*
- 2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
- 3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
- 4: /cgn2_6/ptodata/1/paa/US080_COMB.pep.*
- 5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
- 6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
- 7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
- 8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
- 9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
- 10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
- 11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
- 12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
- 13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
- 14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
- 15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
- 16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
- 17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
- 18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
- 19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
- 20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
- 21: /cgn2_6/ptodata/1/paa/US097_COMB.pep.*
- 22: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
- 23: /cgn2_6/ptodata/1/paa/US099_COMB.pep.*
- 24: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
- 25: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
- 26: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
- 27: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1048	100.0	198	19 US-09-575-580B-4	Sequence 4, Appli
2	1048	100.0	198	21 US-09-782-953-3	Sequence 3, Appli
3	1048	100.0	198	21 US-09-782-953-4	Sequence 4, Appli
4	1002.5	95.7	197	21 US-09-782-953-15	Sequence 15, Appl
5	1002.5	95.7	197	21 US-09-782-953-16	Sequence 16, Appl
6	1002.5	95.7	252	25 US-10-104-047-2892	Sequence 2892, Ap

7	920.5	87.8	198	21	US-09-782-953-6	Sequence 6, Appli
8	920.5	87.8	198	21	US-09-782-953-7	Sequence 7, Appli
9	884.5	84.4	197	19	US-09-575-580B-8	Sequence 8, Appli
10	872.5	83.3	197	21	US-09-782-953-12	Sequence 12, Appli
11	872.5	83.3	197	21	US-09-782-953-13	Sequence 13, Appli
12	872.5	83.3	197	26	US-10-247-671-174	Sequence 174, App
13	872.5	83.3	197	27	US-09-575-580B-7	Sequence 7, Appli
14	843.5	80.5	170	19	US-09-575-580B-7	Sequence 7, Appli
15	620.5	59.2	234	24	US-10-030-613-1	Sequence 1, Appli
16	620.5	59.2	234	27	US-09-142-678-1	Sequence 1, Appli
17	620.5	59.2	243	25	US-10-104-047-2216	Sequence 2216, Ap
18	620.5	59.2	255	20	US-09-614-474-2	Sequence 2, Appli
19	613.5	58.5	197	21	US-09-782-953-9	Sequence 9, Appli
20	613.5	58.5	197	21	US-09-782-953-10	Sequence 10, Appli
21	609.5	58.2	192	19	US-09-575-580B-6	Sequence 6, Appli
22	609.5	58.2	197	1	PCT-US02-17382-132	Sequence 132, App
23	606.5	57.9	242	19	US-09-575-580B-5	Sequence 5, Appli
24	596.5	56.9	192	20	US-09-614-474-10	Sequence 10, Appli
25	596.5	56.9	192	21	US-09-782-953-18	Sequence 18, Appli
26	596.5	56.9	192	21	US-09-782-953-19	Sequence 19, Appli
27	596	56.9	241	1	PCT-US02-17382-133	Sequence 133, App
28	596	56.9	241	20	US-09-614-474-11	Sequence 11, Appli
29	596	56.9	241	21	US-09-782-953-21	Sequence 21, Appli
30	596	56.9	241	21	US-09-782-953-22	Sequence 22, Appli
31	595	56.8	212	21	US-09-782-953-24	Sequence 24, Appli
32	595	56.8	212	21	US-09-782-953-25	Sequence 25, Appli
33	583.5	55.7	236	19	US-09-575-580B-24	Sequence 24, Appli
34	497	47.4	142	1	PCT-US00-05918-790	Sequence 790, App
35	497	47.4	142	23	US-09-525-302-790	Sequence 790, App
36	432	41.2	142	21	US-09-758-472-6466	Sequence 6466, App
37	432	41.2	142	26	US-10-235-926-6466	Sequence 6466, App
38	379.5	36.2	292	20	US-09-614-150-41193	Sequence 41193, A
39	379.5	36.2	292	20	US-09-619-049-1440	Sequence 1440, App
40	379.5	36.2	292	27	US-60-167-324-718	Sequence 718, App
41	379.5	36.2	292	27	US-60-171-627-2109	Sequence 2109, App
42	379.5	36.2	292	27	US-60-173-386-680	Sequence 680, App
43	379.5	36.2	292	27	US-60-175-871-764	Sequence 764, App
44	379.5	36.2	292	27	US-60-184-775-692	Sequence 692, App
45	379.5	36.2	292	27	US-60-191-637-40818	Sequence 40818, A

ALIGNMENTS

RESULT 1
US-09-575-580B-4
; Sequence 4, Application US/09575580B
; GENERAL INFORMATION:
; APPLICANT: McKeon, F.
; APPLICANT: Kayako, K.
; APPLICANT: Rycom, S.
; TITLE OF INVENTION: CALCIPRESSINS: ENDOGENOUS INHIBITORS OF CALCINEURIN
; FILE OF INVENTION: USES AND REAGENTS RELATED THERETO
; FILE REFERENCE: HMV-048.01
; CURRENT APPLICATION NUMBER: US/09/575.580B
; CURRENT FILING DATE: 2000-05-22
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-575-580B-4

Query Match 100.0%; Score 1048; DB 19; Length 198;
Best Local Similarity 100.0%; Pred. No. 7.2e-96;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 MEEVDQLPSATIACHLDPFVVDGLCRACKFESLFTYDKDTTFQYFKSKRVIRFNSN 60
Db 1 MEEVDQLPSATIACHLDPFVVDGLCRACKFESLFTYDKDTTFQYFKSKRVIRFNSN 60
Oy 61 PLSAADARLRLHKTFLGKMKLYFAQTLLHTGSSHLAPPNPDKQFLISPPASPPVGWQV 120

```

Db 61 PLSAADARLRLHKTFFLGKEMKLYFAQTLLHIGSSHLAPPNDKQFLISPPASPPVGMKQV 120
QY 121 EDATPVINVDLLYAISKLGPGKYEELHAATDPTPSVVVHVCSDDQENEEBEEEMERMRKP 180
Db 121 EDATPVINVDLLYAISKLGPGKYEELHAATDPTPSVVVHVCSDDQENEEBEEEMERMRKP 180
QY 181 KPKIIOTRRPEYTPPIHLS 198
Db 181 KPKIIOTRRPEYTPPIHLS 198

```

RESULT 2
US-09-782-953-3
Sequence 3, Application US/09782953

```

; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674P21
; CURRENT APPLICATION NUMBER: US/09/782,953
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 3
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-782-953-3

```

Query Match 100.0%; Score 1048; DB 21; Length 198;
Best Local Similarity 100.0%; Pred. No. 7.2e-96;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MEEVDLQDLP SATIACHLDPRVFDGLCRAKFESLFRITDKDTTFQYFKSFKRVIRINFSN 60
Db 1 MEEVDLQDLP SATIACHLDPRVFDGLCRAKFESLFRITDKDTTFQYFKSFKRVIRINFSN 60
QY 61 PLSAADARLRLHKTFFLGKEMKLYFAQTLLHIGSSHLAPPNDKQFLISPPASPPVGMKQV 120
Db 61 PLSAADARLRLHKTFFLGKEMKLYFAQTLLHIGSSHLAPPNDKQFLISPPASPPVGMKQV 120
QY 121 EDATPVINVDLLYAISKLGPGKYEELHAATDPTPSVVVHVCSDDQENEEBEEEMERMRKP 180
Db 121 EDATPVINVDLLYAISKLGPGKYEELHAATDPTPSVVVHVCSDDQENEEBEEEMERMRKP 180
QY 181 KPKIIOTRRPEYTPPIHLS 198
Db 181 KPKIIOTRRPEYTPPIHLS 198

```

RESULT 3
US-09-782-953-4
Sequence 4, Application US/09782953
GENERAL INFORMATION:
APPLICANT: WILLIAMS, R. SANDERS
APPLICANT: ROTHERMEL, BEVERLY
TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
FILE REFERENCE: UTSD:674P21
CURRENT APPLICATION NUMBER: US/09/782,953
CURRENT FILING DATE: 2001-02-13
PRIOR APPLICATION NUMBER: 60/216,601
PRIOR FILING DATE: 2000-07-07
NUMBER OF SEQ ID NOS: 27
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO: 4
LENGTH: 198
TYPE: PRT
ORGANISM: Mus musculus

US-09-782-953-4

Query Match 100.0%; Score 1048; DB 21; Length 198;
Best Local Similarity 100.0%; Pred. No. 7.2e-96;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MEEVDLQDLP SATIACHLDPRVFDGLCRAKFESLFRITDKDTTFQYFKSFKRVIRINFSN 60
Db 1 MEEVDLQDLP SATIACHLDPRVFDGLCRAKFESLFRITDKDTTFQYFKSFKRVIRINFSN 60
QY 61 PLSAADARLRLHKTFFLGKEMKLYFAQTLLHIGSSHLAPPNDKQFLISPPASPPVGMKQV 120
Db 61 PLSAADARLRLHKTFFLGKEMKLYFAQTLLHIGSSHLAPPNDKQFLISPPASPPVGMKQV 120
QY 121 EDATPVINVDLLYAISKLGPGKYEELHAATDPTPSVVVHVCSDDQENEEBEEEMERMRKP 180
Db 121 EDATPVINVDLLYAISKLGPGKYEELHAATDPTPSVVVHVCSDDQENEEBEEEMERMRKP 180
QY 181 KPKIIOTRRPEYTPPIHLS 198
Db 181 KPKIIOTRRPEYTPPIHLS 198

```

RESULT 4
US-09-782-953-15
Sequence 15, Application US/09782953

```

; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674P21
; CURRENT APPLICATION NUMBER: US/09/782,953
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 15
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-15

```

Query Match 95.7%; Score 1002.5; DB 21; Length 197;
Best Local Similarity 96.5%; Pred. No. 2.5e-91;
Matches 191; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

```

QY 1 MEEVDLQDLP SATIACHLDPRVFDGLCRAKFESLFRITDKDTTFQYFKSFKRVIRINFSN 60
Db 1 MEEVDLQDLP SATIACHLDPRVFDGLCRAKFESLFRITDKDTTFQYFKSFKRVIRINFSN 60
QY 61 PLSAADARLRLHKTFFLGKEMKLYFAQTLLHIGSSHLAPPNDKQFLISPPASPPVGMKQV 120
Db 61 PLSAADARLRLHKTFFLGKEMKLYFAQTLLHIGSSHLAPPNDKQFLISPPASPPVGMKQV 120
QY 121 EDATPVINVDLLYAISKLGPGKYEELHAATDPTPSVVVHVCSDDQENEEBEEEMERMRKP 180
Db 121 EDATPVINVDLLYAISKLGPGKYEELHAATDPTPSVVVHVCSDDQENEEBEEEMERMRKP 180
QY 181 KPKIIOTRRPEYTPPIHLS 198
Db 181 KPKIIOTRRPEYTPPIHLS 198

```

RESULT 5
US-09-782-953-16
Sequence 16, Application US/09782953
GENERAL INFORMATION:
APPLICANT: WILLIAMS, R. SANDERS
APPLICANT: ROTHERMEL, BEVERLY
TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
FILE REFERENCE: UTSD:674P21
CURRENT APPLICATION NUMBER: US/09/782,953
CURRENT FILING DATE: 2001-02-13
PRIOR APPLICATION NUMBER: 60/216,601
PRIOR FILING DATE: 2000-07-07
NUMBER OF SEQ ID NOS: 27
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO: 16
LENGTH: 198
TYPE: PRT
ORGANISM: Mus musculus

```
; FILE REFERENCE: UTSD:674P21
; CURRENT APPLICATION NUMBER: US/09/782,953
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 16
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-16

Query Match      95.7%; Score 1002.5; DB 21; Length 197;
Best Local Similarity 96.5%; Pred. No. 2.5e-91;
Matches 191; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

QY 1 MEEVDLQDLPSATIACHLDPRVFDGLCRAKFESLRTYDKDTTFOYKSKFKVRINFNS 60
Db 1 MEEVDLQDLPSATIACHLDPRVFDGLCRAKFESLRTYDKDTTFOYKSKFKVRINFNS 60

QY 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHGSSHLAPPNDPKQFLISPPASPPVGWQV 120
Db 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHGSSHLAPPNDPKQFLISPPASPPVGWQV 120

QY 121 EDATPVINDLLYAIKSGEGEYELHAATDTPSPVVHVCSQDQNEEEEMERMRP 180
Db 121 EDATPVINDLLYAIKSGEGEYELHAATDTPSPVVHVCSQDQNEEEEMERMRP 179

QY 181 KPILQTRRPEYTPHLS 198
Db 180 KPILQTRRPEYTPHLS 197

RESULT 6
US-10-104-047-2892
; Sequence 2892, Application US/10104047
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cdna
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2892
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-2892

Query Match      95.7%; Score 1002.5; DB 25; Length 252;
Best Local Similarity 96.5%; Pred. No. 3.6e-91;
Matches 191; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

QY 1 MEEVDLQDLPSATIACHLDPRVFDGLCRAKFESLRTYDKDTTFOYKSKFKVRINFNS 60
Db 56 MEEVDLQDLPSATIACHLDPRVFDGLCRAKFESLRTYDKDTTFOYKSKFKVRINFNS 115

QY 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHGSSHLAPPNDPKQFLISPPASPPVGWQV 120
Db 116 PLSAADARLRLHKTFLGKEMKLYFAQTLHGSSHLAPPNDPKQFLISPPASPPVGWQV 175

QY 121 EDATPVINDLLYAIKSGEGEYELHAATDTPSPVVHVCSQDQNEEEEMERMRP 180
Db 176 EDATPVINDLLYAIKSGEGEYELHAATDTPSPVVHVCSQDQNEEEEMERMRP 234

QY 181 KPILQTRRPEYTPHLS 198
Db 235 KPILQTRRPEYTPHLS 252
```

```
RESULT 7
US-09-782-953-6
; Sequence 6, Application US/09782953
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674P21
; CURRENT APPLICATION NUMBER: US/09/782,953
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 6
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-782-953-6

Query Match      87.8%; Score 920.5; DB 21; Length 198;
Best Local Similarity 89.9%; Pred. No. 4e-83;
Matches 178; Conservative 5; Mismatches 12; Indels 3; Gaps 1;

QY 4 VDLQDLP---SATIACHLDPRVFDGLCRAKFESLRTYDKDTTFOYKSKFKVRINFNS 60
Db 1 MDRDFSYNFSLIACVANDVDVSESETRAKFESLRTYDKDTTFOYKSKFKVRINFNS 60

QY 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHGSSHLAPPNDPKQFLISPPASPPVGWQV 120
Db 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHGSSHLAPPNDPKQFLISPPASPPVGWQV 120

QY 121 EDATPVINDLLYAIKSGEGEYELHAATDTPSPVVHVCSQDQNEEEEMERMRP 180
Db 121 EDATPVINDLLYAIKSGEGEYELHAATDTPSPVVHVCSQDQNEEEEMERMRP 180

QY 181 KPILQTRRPEYTPHLS 198
Db 181 KPILQTRRPEYTPHLS 198

RESULT 8
US-09-782-953-7
; Sequence 7, Application US/09782953
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674P21
; CURRENT APPLICATION NUMBER: US/09/782,953
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 7
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-782-953-7

Query Match      87.8%; Score 920.5; DB 21; Length 198;
Best Local Similarity 89.9%; Pred. No. 4e-83;
Matches 178; Conservative 5; Mismatches 12; Indels 3; Gaps 1;

QY 4 VDLQDLP---SATIACHLDPRVFDGLCRAKFESLRTYDKDTTFOYKSKFKVRINFNS 60
Db 1 MDRDFSYNFSLIACVANDVDVSESETRAKFESLRTYDKDTTFOYKSKFKVRINFNS 60

QY 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHGSSHLAPPNDPKQFLISPPASPPVGWQV 120
```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 11:39:24 ; Search time 11.6863 Seconds
(without alignments)
1118.077 Million cell updates/sec

Title: US-09-782-953-3
Perfect score: 1048
Sequence: 1 MEEVDLQPLSATIACHLDP.....RPKPIIOTRRPEYPIHLS 198

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 233758 seqs, 65991041 residues

Total number of hits satisfying chosen parameters: 233758

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents AA New.*
1: /cgn2_6/prodata1/paa/PCT_NEW_COMB.pep.*
2: /cgn2_6/prodata1/paa/US06_NEW_COMB.pep.*
3: /cgn2_6/prodata1/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/prodata1/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/prodata1/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/prodata1/paa/US10_NEW_COMB.pep.*
7: /cgn2_6/prodata1/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	999.5	95.4	197	5	US-09-724-676-61352
2	999.5	95.4	197	5	US-09-724-676-61352
3	872.5	83.3	197	5	US-09-724-676-61358
4	872.5	83.3	197	5	US-09-724-676-61359
5	872.5	83.3	197	5	US-09-724-676-61367
6	872.5	83.3	197	5	US-09-724-676-61358
7	872.5	83.3	197	5	US-09-724-676-61359
8	872.5	83.3	197	5	US-09-724-676-61367
9	848	80.9	200	5	US-09-724-676-61356
10	848	80.9	200	5	US-09-724-676-61356
11	847.5	80.9	171	5	US-09-724-676-61354
12	847.5	80.9	171	5	US-09-724-676-61354
13	731	69.8	154	5	US-09-724-676-61353
14	731	69.8	154	5	US-09-724-676-61353
15	620.5	59.2	255	6	US-10-290-438-2
16	607	57.9	205	5	US-09-724-676-61360
17	607	57.9	205	5	US-09-724-676-61360
18	604	57.6	154	5	US-09-724-676-61361
19	604	57.6	154	5	US-09-724-676-61361
20	603.5	57.6	162	5	US-09-724-676-61351
21	603.5	57.6	162	5	US-09-724-676-61351
22	596.5	56.9	192	6	US-10-290-438-10
23	596	56.9	241	6	US-10-290-438-11
24	580	55.3	156	5	US-09-724-676-61355
25	580	55.3	156	5	US-09-724-676-61355
26	579.5	55.3	157	5	US-09-724-676-61357

ALIGNMENTS

RESULT 1

US-09-724-676-61352
; Sequence 61352, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 61352
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-61352

Query Match 95.4%; Score 999.5; DB 5; Length 197;
Best Local Similarity 96.0%; Pred. No. 2.1e-94;
Matches 190; Conservative 4; Mismatches 3; Indels 1; Gaps 1;

Qy	1	MEEVDLQPLSATIACHLDP	RVVDCGLCRAPESLRTYDKDITTFQYKSFKEVRINFSN	60
Db	1	MEEVDLQPLSATIACHLDP	RVVDCGLCRAPESLRTYDKDITTFQYKSFKEVRINFSN	60
Qy	61	PLSAADARLRLHKTEFLGKEMKLYFAQTLLHIGSSHLAPPNDKOFLLISPPASPPVGWKQV	120	
Db	61	PLSAADARLRLHKTEFLGKEMKLYFAQTLLHIGSSHLAPPNDKOFLLISPPASPPVGWKQV	120	
Qy	121	EDATPVINDLLYAIKSLGPGKYEYELHAATDTPPSVVHVCSDDQNEEEEMERMRP	180	
Db	121	EDATPVINDLLYAIKSLGPGKYEYELHAATDTPPSVVHVCSDDQNEEEEMERMRP	179	
Qy	181	KPKIIOTRRPEYPIHLS	198	
Db	180	KPKIIOTRRPEYPIHLS	197	

RESULT 2

US-09-724-676A-61352
; Sequence 61352, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2

```
; SEQ ID NO 61352
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-61352
```

```
Query Match          95.4%; Score 999.5; DB 5; Length 197;
Best Local Similarity 96.0%; Pred. No. 2.1e-94;
Matches 190; Conservative 4; Mismatches 3; Indels 1; Gaps 1;
```

```
QY 1 MEEVLOLPSTATICHLDPRVFDGLCRAKFESLFRITYDKDTTFQYFKSFKRVINFSN 60
DB 1 MEEVLEDPSTATICHLDPRVFDGLCRAKFESLFRITYDKDTTFQYFKSFKRVINFSN 60
QY 61 PLASADARLRHKTEFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPPVGMKQV 120
DB 61 PPSAADARLRQHKTEFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPPVGMKQV 120
QY 121 EDATFVINDDLVAISKLGPGGKYEELHAATDPTPSVVHVHVCESDQENESEEMERMRKP 180
DB 121 EDATFVINDDLVAISKLGPGGKYEELHAATDPTPSVVHVHVCESDQENESEEMERMRKP 179
QY 181 KPRTIQTRRPEYPIHLS 198
DB 180 KPRTIQTRRPEYPIHLS 197
```

```
RESULT 3
US-09-724-676-61358
; Sequence 61358, Application US/09724676
; GENERAL INFORMATION:
```

```
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 61358
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-61358
```

```
Query Match          83.3%; Score 872.5; DB 5; Length 197;
Best Local Similarity 89.4%; Pred. No. 1.9e-81;
Matches 168; Conservative 7; Mismatches 12; Indels 1; Gaps 1;
```

```
QY 11 SATIACHLDPRVFDGLCRAKFESLFRITYDKDTTFQYFKSFKRVINFSNPLSADARLR 70
DB 11 SSLIACVANSDFSESETRAKFESLFRITYDKDTTFQYFKSFKRVINFSNPSAADARLRQ 70
QY 71 LKTEFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPPVGMKQVEDATPVINYD 130
DB 71 LKTEFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPPVGMKQVEDATPVINYD 130
QY 131 LLVAISKLGPGGKYEELHAATDPTPSVVHVHVCESDQENESEEMERMRKPRTKIQTTRP 190
DB 131 LLVAISKLGPGGKYEELHAATDPTPSVVHVHVCESDQENESEEMERMRKPRTKIQTTRP 189
QY 191 EYTPHLS 198
DB 190 EYTPHLS 197
```

```
RESULT 4
US-09-724-676-61359
; Sequence 61359, Application US/09724676
; GENERAL INFORMATION:
```

```
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
```

```
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 61359
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-61359
```

```
Query Match          83.3%; Score 872.5; DB 5; Length 197;
Best Local Similarity 89.4%; Pred. No. 1.9e-81;
Matches 168; Conservative 7; Mismatches 12; Indels 1; Gaps 1;
```

```
QY 11 SATIACHLDPRVFDGLCRAKFESLFRITYDKDTTFQYFKSFKRVINFSNPLSADARLR 70
DB 11 SSLIACVANSDFSESETRAKFESLFRITYDKDTTFQYFKSFKRVINFSNPSAADARLRQ 70
QY 71 LKTEFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPPVGMKQVEDATPVINYD 130
DB 71 LKTEFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPPVGMKQVEDATPVINYD 130
QY 131 LLVAISKLGPGGKYEELHAATDPTPSVVHVHVCESDQENESEEMERMRKPRTKIQTTRP 190
DB 131 LLVAISKLGPGGKYEELHAATDPTPSVVHVHVCESDQENESEEMERMRKPRTKIQTTRP 189
QY 191 EYTPHLS 198
DB 190 EYTPHLS 197
```

```
RESULT 5
US-09-724-676-61367
; Sequence 61367, Application US/09724676
; GENERAL INFORMATION:
```

```
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 61367
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-61367
```

```
Query Match          83.3%; Score 872.5; DB 5; Length 197;
Best Local Similarity 89.4%; Pred. No. 1.9e-81;
Matches 168; Conservative 7; Mismatches 12; Indels 1; Gaps 1;
```

```
QY 11 SATIACHLDPRVFDGLCRAKFESLFRITYDKDTTFQYFKSFKRVINFSNPLSADARLR 70
DB 11 SSLIACVANSDFSESETRAKFESLFRITYDKDTTFQYFKSFKRVINFSNPSAADARLRQ 70
QY 71 LKTEFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPPVGMKQVEDATPVINYD 130
DB 71 LKTEFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPPVGMKQVEDATPVINYD 130
QY 131 LLVAISKLGPGGKYEELHAATDPTPSVVHVHVCESDQENESEEMERMRKPRTKIQTTRP 190
DB 131 LLVAISKLGPGGKYEELHAATDPTPSVVHVHVCESDQENESEEMERMRKPRTKIQTTRP 189
QY 191 EYTPHLS 198
DB 190 EYTPHLS 197
```

```
RESULT 6
US-09-724-676A-61358
; Sequence 61358, Application US/09724676A
; GENERAL INFORMATION:
```

```
; APPLICANT: Compugen LTD
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 11:37:41 ; Search time 11.3524 Seconds
(without alignments)
513.170 Million cell updates/sec

Title: US-09-782-953-7
Perfect score: 1043
Sequence: 1 MPRDFSYNFSLLIACVAND.....RPKPKIIQTRPEYTPHLS 198

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/prodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/prodata/1/iaa/PCUS_COMB.pep:*
6: /cgn2_6/prodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	835.5	80.1	171	2	US-08-665-040-2
2	84.5	8.1	719	4	US-09-641-741-28
3	82	7.9	284	5	PC-TUS94-09752-4
4	82	7.9	1128	4	US-09-641-741-30
5	82	7.9	1128	4	US-09-060-482-8
6	81	7.8	1128	1	US-08-111-939-2
7	81	7.8	2237	1	US-08-455-543A-48
8	81	7.8	2237	2	US-08-223-305C-48
9	81	7.8	2237	4	US-09-268-163-8
10	81	7.8	2339	1	US-08-455-543A-47
11	81	7.8	2339	2	US-08-223-305C-47
12	81	7.8	2339	4	US-09-268-163-6
13	81	7.8	2343	4	US-09-268-163-4
14	80	7.7	370	2	US-08-231-342-6
15	79.5	7.6	434	2	US-08-815-718-3
16	79.5	7.6	559	2	US-08-884-072-6
17	79.5	7.6	559	4	US-09-213-168-6
18	78.5	7.5	201	2	US-08-716-317-1
19	78.5	7.5	547	4	US-09-877-730-26
20	78.5	7.5	624	4	US-09-877-730-24
21	78.5	7.5	628	4	US-09-877-730-30
22	78.5	7.5	712	4	US-09-877-730-22
23	78.5	7.5	793	4	US-09-877-730-28
24	78.5	7.5	826	4	US-09-877-730-16
25	78.5	7.5	904	4	US-09-877-730-6
26	78.5	7.5	907	4	US-09-877-730-20
27	78.5	7.5	985	4	US-09-877-730-10

28	78.5	7.5	991	4	US-09-877-730-12
29	78.5	7.5	1069	4	US-09-877-730-2
30	78.5	7.5	1072	4	US-09-877-730-18
31	78.5	7.5	1150	4	US-09-877-730-8
32	78.5	7.5	1706	2	US-08-459-568-2
33	78.5	7.5	1706	2	US-08-399-411-2
34	78.5	7.5	1706	3	US-08-516-859A-2
35	78.5	7.5	1706	4	US-09-586-472-2
36	78.5	7.5	1706	4	US-09-528-706-2
37	77	7.4	522	6	RE34606-6
38	77	7.4	1912	4	US-08-913-832A-2
39	77	7.4	1912	4	US-09-249-181A-2
40	77	7.4	2337	3	US-08-713-118-2
41	77	7.4	2337	4	US-09-452-007-2
42	75	7.2	362	2	US-08-985-090-5
43	75	7.2	362	3	US-09-165-543-32
44	75	7.2	1158	4	US-09-060-482-2
45	74	7.1	445	3	US-09-165-543-5

ALIGNMENTS

RESULT 1

US-08-665-040-2
; Sequence 2, Application US/08665040
; Patent No. 5869318
; GENERAL INFORMATION:
; APPLICANT: ESTIVILL PALLEJA, XAVIER
; APPLICANT: FUENTES, JUAN JOSE
; APPLICANT: PRITCHARD, MELANIE
; TITLE OF INVENTION: A NEW GENE SEQUENCE OF THE
; TITLE OF INVENTION: DOWN SYNDROM CRITICAL REGION OF HUMAN CHROMOSOME 21,
; TITLE OF INVENTION: IDENTIFIED BY A NEW "ALU-SPlicing PCR" TECHNIQUE,
; TITLE OF INVENTION: CODING FOR A PROLINE-RICH PROTEIN (DSCR1) HIGHLY
; TITLE OF INVENTION: EXPRESSED IN FOETAL BRAIN AND IN HEART AND METHOD
; TITLE OF INVENTION: FOR CHARACTERIZING IT.
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LADAS & PARRY
; STREET: 26 WEST 61ST STREET
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10023

COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORD PERFECT 5.1 FOR DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/665,040
; FILING DATE: JUNE 7, 1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: ES P9501140
; FILING DATE: JUNE 7, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: JANET I. CORD
; REGISTRATION NUMBER: 33,778
; REFERENCE/DOCKET NUMBER: U010815-9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 708-1800
; TELEFAX: (212) 246-8959
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 171 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-665-040-2

Query Match 80.1%; Score 835.5; DB 2; Length 171;
 Best Local Similarity 94.7%; Pred. No. 4,4e-87;
 Matches 160; Conservative 5; Mismatches 3; Indels 1; Gaps 1;

QY 30 AKFESLFRTYDQTPQYFKSKRYRINFSNPLSADARLRLHKEFLGKEMKLYFAOTL 89
 DB 4 AKFESLFRTYDQTPQYFKSKRYRINFSNPLSADARLRLHKEFLGKEMKLYFAOTL 63
 QY 90 HIGSSHLAPNDKQPLSPSPVGVKQVADATPVINYDLVAISKLGPEKYEELHAA 149
 DB 64 HIGSSHLAPNDKQPLSPSPVGVKQVADATPVINYDLVAISKLGPEKYEELHAA 123
 QY 150 TDPTPSVVHVCESDQENEEEMERMKRPKPIIOTRRPEYTPHLS 198
 DB 124 TDPTPSVVHVCESDQENEEEMERMKRPKPIIOTRRPEYTPHLS 171

RESULT 2
 US-09-641-741-28
 Sequence 28, Application US/09641741

GENERAL INFORMATION:
 APPLICANT: Kerry E. Quinn
 TITLE OF INVENTION: Aortic Carboxypeptidase-like Proteins and Nucleic Acids
 TITLE OF INVENTION: encoding Same
 FILE REFERENCE: 15966-581
 CURRENT APPLICATION NUMBER: US/09/641,741
 CURRENT FILING DATE: 2000-08-18
 PRIOR APPLICATION NUMBER: 60/159,613
 PRIOR FILING DATE: 1999-10-14
 PRIOR APPLICATION NUMBER: 60/175,534
 PRIOR FILING DATE: 2000-01-12
 PRIOR APPLICATION NUMBER: 60/224,086
 PRIOR FILING DATE: 2000-08-09
 NUMBER OF SEQ ID NOS: 32
 SOFTWARE: Patent In Ver. 2.0
 SEQ ID NO 28
 LENGTH: 719
 TYPE: PRT
 ORGANISM: Mus musculus
 US-09-641-741-28

Query Match 8.1%; Score 84.5; DB 4; Length 719;
 Best Local Similarity 21.8%; Pred. No. 0.71;
 Matches 31; Conservative 27; Mismatches 37; Indels 47; Gaps 7;

QY 54 VRINFSNPLSADARLRLHKEFLGKEMKLYFAOTLHIGSSHLAPNDKQPLSP 110
 DB 585 LRVDPSRPMTPOORRMQRLQYLRMRBQMR--RLNLTAGPATSPTP--ALMP 637
 QY 111 ASP-----PVGMK--QVEDATPVINYDLVAISKLGPEKYEELHAA 153
 DB 638 PSPTPAITLRFWEVLPTTIGMESEETETVTVT-----EFETRYGTD-- 681
 QY 154 PSVVHVCESDQENEEEMERMKRPKPIIOTRRPEYTPHLS 175
 DB 682 -----LEVEELEEEEEEEMD 699

RESULT 3
 PCT-US94-09752-4
 Sequence 4, Application PC/TUS9409752
 GENERAL INFORMATION:
 APPLICANT: David S. Strayer and Arinash Chander
 TITLE OF INVENTION: Compositions and Methods for
 TITLE OF INVENTION: Targeting Cells and Modulating Pulmonary Surfactant Secretion
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Jane Massey Licata, Esq.
 STREET: 210 Lake Drive East, Suite 201
 CITY: Cherry Hill
 STATE: NJ

COUNTRY: USA
 ZIP: 08002
 COMPUTER READABLE FORM:
 MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb
 MEDIUM TYPE: STORAGE
 COMPUTER: IBM PS/2
 OPERATING SYSTEM: PC-DOS
 SOFTWARE: WORDPERFECT 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US94/09752
 FILING DATE: Herewith
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/176,218
 FILING DATE: December 30, 1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/114,951
 FILING DATE: August 31, 1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Jane Massey Licata
 REGISTRATION NUMBER: 32,257
 REFERENCE/DOCKET NUMBER: JEFF-0042
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (609) 779-8488
 TELEFAX: (609) 779-8488
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 284
 TYPE: amino acid
 TOPOLOGY: linear
 PCT-US94-09752-4

Query Match 7.9%; Score 82; DB 5; Length 284;
 Best Local Similarity 24.5%; Pred. No. 0.35;
 Matches 39; Conservative 23; Mismatches 57; Indels 40; Gaps 7;

QY 34 SLFRYDDQTPQYFKSKRYRINFSNPLSADARLRLHKEFLGKEMKLYFAOTLHIGS 93
 DB 10 SLFRSGRQTLPA-----SOVMRYAD--LYAASFVN--LTVPPSYLFR 50
 QY 94 SHLAPPNP-----DKQPLSPAS--PVGMKQVEDATPVINYDLVAISKLGPE 142
 DB 51 AAVLPHSHTVHTHVDLINESPLATNRKTSVDFK-----TDYRHHQLTSSISIKRPN 106
 QY 143 KYELHAATDPTPSVVHVCESDQENEEEMERMKRPK 181
 DB 107 LFPL-----APQIITHCHDEDDDEEEBEECGKPK 139

RESULT 4
 US-09-641-741-30
 Sequence 30, Application US/09641741
 Patent No. 6420155
 GENERAL INFORMATION:
 APPLICANT: Kerry E. Quinn
 TITLE OF INVENTION: Aortic Carboxypeptidase-like Proteins and Nucleic Acids
 TITLE OF INVENTION: encoding Same
 FILE REFERENCE: 15966-581
 CURRENT APPLICATION NUMBER: US/09/641,741
 CURRENT FILING DATE: 2000-08-18
 PRIOR APPLICATION NUMBER: 60/159,613
 PRIOR FILING DATE: 1999-10-14
 PRIOR APPLICATION NUMBER: 60/175,534
 PRIOR FILING DATE: 2000-01-12
 PRIOR APPLICATION NUMBER: 60/224,086
 PRIOR FILING DATE: 2000-08-09
 NUMBER OF SEQ ID NOS: 32
 SOFTWARE: Patent In Ver. 2.0
 SEQ ID NO 30
 LENGTH: 1128
 TYPE: PRT
 ORGANISM: Mus musculus

US-09-641-741-30

Query Match 7.9%; Score 82; DB 4; Length 1128;
Best Local Similarity 20.8%; Pred. No. 2.6;
Matches 33; Conservative 29; Mismatches 37; Indels 60; Gaps 8;

QY 50 SFKRV-----INFNPLSADARLRHKTFF---LGKMKLYFAOTLHIGS 93
DB 977 NWKRIEILANGNRPIGVDPSPMTPOQRMQORRLQYRLRMREQMQL---RRLNSTA 1033
QY 94 SHLAPPNDKQFLISPPASP-----PVGWK--QVEDATPVINYDLYAIS 136
DB 1034 GPATSPTP-----ALMPPSPPTAITLRPWVLPPTTAGWSESETETYTEVVT----- 1081
QY 137 KLGPGEKVELHAATDPTSPVVHVCSQDNEEEEEEME 175
DB 1082 -----EFETEYGTD-----LEVEIEEEEEEEEMD 1108

RESULT 5

US-09-060-482-8
; Sequence 8, Application US/09060482
; Patent No. 6468766
; GENERAL INFORMATION:
; APPLICANT: Lee, Mu-En
; APPLICANT: Layne, Matthew D.
; APPLICANT: Yet, Shaw-Fang
; TITLE OF INVENTION: AORTIC CARDOXYPEPTIDASE-LIKE POLYPEPTIDE
; FILE REFERENCE: 05433/036001
; CURRENT APPLICATION NUMBER: US/09/060,482
; CURRENT FILING DATE: 1998-04-15
; EARLIER APPLICATION NUMBER: US 081818,009
; EARLIER FILING DATE: 1997-03-14
; EARLIER APPLICATION NUMBER: US 60\013,439
; EARLIER FILING DATE: 1996-03-15
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 1128
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-060-482-8

Query Match 7.9%; Score 82; DB 4; Length 1128;
Best Local Similarity 20.8%; Pred. No. 2.6;
Matches 33; Conservative 29; Mismatches 37; Indels 60; Gaps 8;

QY 50 SFKRV-----INFNPLSADARLRHKTFF---LGKMKLYFAOTLHIGS 93
DB 977 NWKRIEILANGNRPIGVDPSPMTPOQRMQORRLQYRLRMREQMQL---RRLNSTA 1033
QY 94 SHLAPPNDKQFLISPPASP-----PVGWK--QVEDATPVINYDLYAIS 136
DB 1034 GPATSPTP-----ALMPPSPPTAITLRPWVLPPTTAGWSESETETYTEVVT----- 1081
QY 137 KLGPGEKVELHAATDPTSPVVHVCSQDNEEEEEEME 175
DB 1082 -----EFETEYGTD-----LEVEIEEEEEEEEMD 1108

RESULT 6

US-08-111-939-2
; Sequence 2, Application US/08111939
; Patent No. 5460951
; GENERAL INFORMATION:
; APPLICANT: Kawai, Shinji
; APPLICANT: Takeshita, Sunao
; APPLICANT: Okazaki, Makoto
; APPLICANT: Amann, Egon
; TITLE OF INVENTION: Bone-Related Carboxypeptidase-Like
; TITLE OF INVENTION: Protein and Process for its Production
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESSEE: Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/111,939
FILING DATE: 26-AUG-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 324033/92
FILING DATE: 03-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 230029/92
FILING DATE: 28-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: Forman, David S.
REGISTRATION NUMBER: 33,694
REFERENCE/DOCKET NUMBER: 02481.1321-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4000
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1128 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-111-939-2

Query Match 7.8%; Score 81; DB 1; Length 1128;
Best Local Similarity 20.8%; Pred. No. 3.4;
Matches 33; Conservative 29; Mismatches 37; Indels 60; Gaps 8;

QY 50 SFKRV-----INFNPLSADARLRHKTFF---LGKMKLYFAOTLHIGS 93
DB 977 NWKRIEILANGNRPIGVDPSPMTPOQRMQORRLQYRLRMREQMQL---RRLNSTA 1033
QY 94 SHLAPPNDKQFLISPPASP-----PVGWK--QVEDATPVINYDLYAIS 136
DB 1034 GPATSPTP-----ALMPPSPPTAITLRPWVLPPTTAGWSESETETYTEVVT----- 1081
QY 137 KLGPGEKVELHAATDPTSPVVHVCSQDNEEEEEEME 175
DB 1082 -----EFETEYGTD-----LEVEIEEEEEEEEMD 1108

RESULT 7

US-08-455-543A-48
; Sequence 48, Application US/08455543A
; Patent No. 5792846
; GENERAL INFORMATION:
; APPLICANT: Harpold, Michael
; APPLICANT: Ellis, Steven
; APPLICANT: Williams, Mark
; APPLICANT: Feldman, Daniel
; APPLICANT: McCue, Ann
; APPLICANT: Brenner, Robert
; TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
ADDRESSEE: Brown, Martin, Haller & McClain
STREET: 1660 Union Street
CITY: San Diego
STATE: California

QY 67 -----ARLRH-----KTEFLGKEM-----KLYFAOTLHGSSHLA 97
Db 752 ISIAARQONSARQSVWEORASQLRNLQRASCEALYSEMDPEERLRFATTHL-----805
QY 98 PPNPDKQFLISPP-----ASPPVGWQKVEDATPVINYDLLYALSKLGPGEKVELHA 148
Db 806 --RPDMKTHLDRLVLVVELGRDARGVGGKARPEAAE-----APEGVDPFRHRRHR 855
QY 149 ATDPTPSVVHVHVCESDQE-----NEEEEMERMKRKP 182
Db 856 DKDKTPAA-----GQDRAEAPKAESGEPGAREERPRP 888

RESULT 9
US-09-268-163-8
; Sequence 8, Application US/09268163B
; Patent No. 6353091
; GENERAL INFORMATION:
; APPLICANT: Lipscombe, Diane
; APPLICANT: Schorge, Stephanie
; TITLE OF INVENTION: HUMAN N-TYPE CALCIUM CHANNEL ISOFORM AND USES THEREOF
; FILE REFERENCE: B1055/7000
; CURRENT APPLICATION NUMBER: US/09/268,163B
; CURRENT FILING DATE: 1999-03-12
; EARLIER APPLICATION NUMBER: US 60/077,901
; EARLIER FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 2237
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-268-163-8

Query Match 7.8%; Score 81; DB 4; Length 2237;
Best Local Similarity 21.5%; Pred. No. 9.4;
Matches 47; Conservative 27; Mismatches 75; Indels 70; Gaps 9;

QY 7 SYNFSLLIACVANDVFESETRAKFESLFRITYDKDTTFQYFKSKRVRINFNSNPLSAAD 66
Db 697 NYTLNVFLAIAVDNLANAQELTKDEEMEEAANQKALQAKAEVAEV-----SPMSAAN 751
QY 67 -----ARLRH-----KTEFLGKEM-----KLYFAOTLHGSSHLA 97
Db 752 ISIAARQONSARQSVWEORASQLRNLQRASCEALYSEMDPEERLRFATTHL-----805
QY 98 PPNPDKQFLISPP-----ASPPVGWQKVEDATPVINYDLLYALSKLGPGEKVELHA 148
Db 806 --RPDMKTHLDRLVLVVELGRDARGVGGKARPEAAE-----APEGVDPFRHRRHR 855
QY 149 ATDPTPSVVHVHVCESDQE-----NEEEEMERMKRKP 182
Db 856 DKDKTPAA-----GQDRAEAPKAESGEPGAREERPRP 888

RESULT 10
US-08-455-543A-47
; Sequence 47, Application US/08455543A
; Patent No. 5792846
; GENERAL INFORMATION:
; APPLICANT: Harpold, Michael
; APPLICANT: Ellis, Steven
; APPLICANT: Williams, Mark
; APPLICANT: Feldman, Daniel
; APPLICANT: McCue, Ann
; APPLICANT: Brenner, Robert
; TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street

CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101-2926
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/455,543A
FILING DATE: May 31, 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/223,305
FILING DATE: April 4, 1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/868,354
FILING DATE: April 10, 1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/745,206
FILING DATE: 15-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/620,250
FILING DATE: 30-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/482,384
FILING DATE: 20-FEB-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/603,751
FILING DATE: 04-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US89/01408
FILING DATE: 04-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/176,899
FILING DATE: 04-APR-1988
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 6362-52517
TELEPHONE: (619) 238-0999
TELEFAX: (619) 238-0062
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 2339 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-455-543A-47

Query Match 7.8%; Score 81; DB 1; Length 2339;
Best Local Similarity 21.5%; Pred. No. 10;
Matches 47; Conservative 27; Mismatches 75; Indels 70; Gaps 9;

QY 7 SYNFSLLIACVANDVFESETRAKFESLFRITYDKDTTFQYFKSKRVRINFNSNPLSAAD 66
Db 697 NYTLNVFLAIAVDNLANAQELTKDEEMEEAANQKALQAKAEVAEV-----SPMSAAN 751
QY 67 -----ARLRH-----KTEFLGKEM-----KLYFAOTLHGSSHLA 97
Db 752 ISIAARQONSARQSVWEORASQLRNLQRASCEALYSEMDPEERLRFATTHL-----805
QY 98 PPNPDKQFLISPP-----ASPPVGWQKVEDATPVINYDLLYALSKLGPGEKVELHA 148
Db 806 --RPDMKTHLDRLVLVVELGRDARGVGGKARPEAAE-----APEGVDPFRHRRHR 855
QY 149 ATDPTPSVVHVHVCESDQE-----NEEEEMERMKRKP 182
Db 856 DKDKTPAA-----GQDRAEAPKAESGEPGAREERPRP 888

RESULT 11
 US-08-223-305C-47
 Sequence 47, Application US/08223305C
 Patent No. 5851824
 GENERAL INFORMATION:
 APPLICANT: Harpold, Michael
 APPLICANT: Ellis, Steven
 APPLICANT: Williams, Mark
 APPLICANT: Feldman, Daniel
 APPLICANT: McCue, Ann
 APPLICANT: Brenner, Robert
 TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
 TITLE OF INVENTION: METHODS
 NUMBER OF SEQUENCES: 57
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Brown, Martin, Haller & McClain
 STREET: 1660 Union Street
 CITY: San Diego
 STATE: California
 COUNTRY: USA
 ZIP: 92101-22926
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq Version 1.5
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/223.305C
 FILING DATE: April 4, 1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/868,354
 FILING DATE: April 10, 1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/745,206
 FILING DATE: 15-AUG-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/620,250
 FILING DATE: 30-NOV-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/482,384
 FILING DATE: 20-FEB-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/603,751
 FILING DATE: 04-APR-1989
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: WO PCT/US89/01408
 FILING DATE: 04-APR-1989
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/176,899
 FILING DATE: 04-APR-1988
 ATTORNEY/AGENT INFORMATION:
 NAME: Seidman, Stephanie L.
 REGISTRATION NUMBER: 33,779
 REFERENCE/DOCKET NUMBER: 52516 (P519739)
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619)238-0999
 TELEFAX: (619)238-0062
 INFORMATION FOR SEQ ID NO: 47:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2339 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 FRAGMENT TYPE: internal
 US-08-223-305C-47

Query Match 7.8%; Score 81; DB 2; Length 2339;
 Best Local Similarity 21.5%; Pred. No. 10;
 Matches 47; Conservative 27; Mismatches 75; Indels 70; Gaps 9;
 QY 7 SYNFSLLIACVANDVSESESETRAKFESLFRITYDKDTFOYFKSKRVIRINFSNPLSAAD 66

Db 697 NYTLNVPFLAIVDNLNAOQLTKDEEBEEMERANOKLALQAKAEVAAY-----SPMSAN 751
 QY 67 -----ARLRH-----KTEFLGKEM-----KLYPEAOTLHIGSSHLA 97
 Db 752 ISIAARQONSAAKARSWEQGRASQLRLQNLRASCEALYSEMDPEERLRFATRRHL----- 805
 QY 98 PPNPDQKQFLISPP-----ASPPVGWKQVEDATPVINYDLIYAIKLGQGEKYTELHA 148
 Db 806 --RPDMKTHLDRLPVLVELGRDARGPVGKARPEAAE-----ABEGVDPPIRRHHR 855
 QY 149 ATDPPTSVVHVHVESDOE-----NEEEEMEEMERMKRPPK 182
 Db 856 DKDKTPAA-----GDQDRAPAPKAESEGPARERRPP 888

RESULT 12
 US-09-268-163-6
 Sequence 6, Application US/09268163B
 Patent No. 6353091
 GENERAL INFORMATION:
 APPLICANT: Lipscombe, Diane
 TITLE OF INVENTION: HUMAN N-TYPE CALCIUM CHANNEL ISOFORM AND USES THEREOF
 FILE REFERENCE: B1055/7000
 CURRENT APPLICATION NUMBER: US/09/268,163B
 CURRENT FILING DATE: 1999-03-12
 EARLIER APPLICATION NUMBER: US 60/077,901
 EARLIER FILING DATE: 1998-03-13
 NUMBER OF SEQ ID NOS: 28
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 6
 LENGTH: 2339
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-268-163-6

Query Match 7.8%; Score 81; DB 4; Length 2339;
 Best Local Similarity 21.5%; Pred. No. 10;
 Matches 47; Conservative 27; Mismatches 75; Indels 70; Gaps 9;

QY 7 SYNFSLLIACVANDVSESESETRAKFESLFRITYDKDTFOYFKSKRVIRINFSNPLSAAD 66
 Db 697 NYTLNVPFLAIVDNLNAOQLTKDEEBEEMERANOKLALQAKAEVAAY-----SPMSAN 751
 QY 67 -----ARLRH-----KTEFLGKEM-----KLYPEAOTLHIGSSHLA 97
 Db 752 ISIAARQONSAAKARSWEQGRASQLRLQNLRASCEALYSEMDPEERLRFATRRHL----- 805
 QY 98 PPNPDQKQFLISPP-----ASPPVGWKQVEDATPVINYDLIYAIKLGQGEKYTELHA 148
 Db 806 --RPDMKTHLDRLPVLVELGRDARGPVGKARPEAAE-----ABEGVDPPIRRHHR 855
 QY 149 ATDPPTSVVHVHVESDOE-----NEEEEMEEMERMKRPPK 182
 Db 856 DKDKTPAA-----GDQDRAPAPKAESEGPARERRPP 888

RESULT 13
 US-09-268-163-4
 Sequence 4, Application US/09268163B
 Patent No. 6353091
 GENERAL INFORMATION:
 APPLICANT: Lipscombe, Diane
 TITLE OF INVENTION: HUMAN N-TYPE CALCIUM CHANNEL ISOFORM AND USES THEREOF
 FILE REFERENCE: B1055/7000
 CURRENT APPLICATION NUMBER: US/09/268,163B
 CURRENT FILING DATE: 1999-03-12
 EARLIER APPLICATION NUMBER: US 60/077,901
 EARLIER FILING DATE: 1998-03-13
 NUMBER OF SEQ ID NOS: 28
 SOFTWARE: FastSeq for Windows Version 3.0

Db 10 ILACIALASASEGAFKASDQ--REMTPERLFOHLH-----EVGYAPPSLPQTRRLRV 61
QY 72 HKTEFLGKEMKLYFAQTLHGSSHLAPNPDQFLLISPPASPPVGKQVEDATPVINYD- 130
Db 62 D-----HSTVSLHDPPLFEHQREVQPPSP-----ED-IPVEEDM 96
QY 131 ---LTYAISKLGPGEKYEELHAATDPTPSVVHVACESDQENEBEEMERMKRPKXIQT 187
Db 97 PTFLLPNVVDKAGPAVQEAIPLQKEQPPQVHI-----EKEIDPPAQPOBEIYQK 147
QY 188 RREXT 193
Db 148 EVKPHI 153

Search completed: December 11, 2002, 11:39:15
Job time : 14.3524 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

QM protein - protein search, using sw model

Run on: December 11, 2002, 11:39:24 ; Search time 11.6863 Seconds
(without alignments)
1118.077 Million cell updates/sec

Title: US-09-782-953-7
Perfect score: 1043
Sequence: 1 MDRFDSYNFSLIACVAND.....RPKPKLIQTRRPEYTPHLS 198

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 233758

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA New.*

- 1: /cgn2_6/prodata1/paa/PCT_NEW_COMB.pep.*
- 2: /cgn2_6/prodata1/paa/US06_NEW_COMB.pep.*
- 3: /cgn2_6/prodata1/paa/US07_NEW_COMB.pep.*
- 4: /cgn2_6/prodata1/paa/US08_NEW_COMB.pep.*
- 5: /cgn2_6/prodata1/paa/US09_NEW_COMB.pep.*
- 6: /cgn2_6/prodata1/paa/US10_NEW_COMB.pep.*
- 7: /cgn2_6/prodata1/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	970.5	93.0	197	5	US-09-724-676-61358
2	970.5	93.0	197	5	US-09-724-676-61359
3	970.5	93.0	197	5	US-09-724-676-61367
4	970.5	93.0	197	5	US-09-724-676-61358
5	970.5	93.0	197	5	US-09-724-676-61359
6	970.5	93.0	197	5	US-09-724-676-61367
7	874	83.8	197	5	US-09-724-676-61352
8	874	83.8	197	5	US-09-724-676-61352
9	847.5	81.3	171	5	US-09-724-676-61354
10	847.5	81.3	171	5	US-09-724-676-61354
11	847.5	81.3	200	5	US-09-724-676-61356
12	847.5	81.3	200	5	US-09-724-676-61356
13	705	67.6	205	5	US-09-724-676-61360
14	705	67.6	205	5	US-09-724-676-61360
15	702	67.3	154	5	US-09-724-676-61361
16	702	67.3	154	5	US-09-724-676-61361
17	701.5	67.3	162	5	US-09-724-676-61351
18	701.5	67.3	162	5	US-09-724-676-61351
19	623	59.7	255	6	US-10-290-438-2
20	615.5	59.0	192	6	US-10-290-438-10
21	605.5	58.1	154	5	US-09-724-676-61353
22	605.5	58.1	154	5	US-09-724-676-61353
23	597.5	57.3	184	5	US-09-724-676-63940
24	597.5	57.3	184	5	US-09-724-676-63940
25	592.5	56.8	241	6	US-10-290-438-11
26	580	55.6	156	5	US-09-724-676-61355

27	580	55.6	156	5	US-09-724-676A-61355	Sequence 61355, A
28	579	55.5	157	5	US-09-724-676A-61357	Sequence 61357, A
29	579	55.5	157	5	US-09-724-676A-61357	Sequence 61357, A
30	537	51.5	111	5	US-09-513-999C-5849	Sequence 5849, Ap
31	415.5	39.8	96	5	US-09-724-676-61363	Sequence 61363, A
32	415.5	39.8	96	5	US-09-724-676-61366	Sequence 61366, A
33	415.5	39.8	96	5	US-09-724-676A-61363	Sequence 61363, A
34	415.5	39.8	96	5	US-09-724-676A-61366	Sequence 61366, A
35	284	27.2	58	6	US-10-203-138A-12463	Sequence 12463, A
36	271.5	26.0	56	6	US-10-203-138A-11146	Sequence 11146, A
37	150	14.4	104	5	US-09-724-676-61364	Sequence 61364, A
38	150	14.4	104	5	US-09-724-676A-61364	Sequence 61364, A
39	146.5	14.0	61	5	US-09-724-676A-61365	Sequence 61365, A
40	146.5	14.0	61	5	US-09-724-676A-61365	Sequence 61365, A
41	127.5	12.2	80	5	US-09-724-676-61362	Sequence 61362, A
42	127.5	12.2	80	5	US-09-724-676A-61362	Sequence 61362, A
43	81	7.8	396	5	US-09-724-676-68208	Sequence 68208, A
44	81	7.8	396	5	US-09-724-676A-68208	Sequence 68208, A
45	81	7.8	434	5	US-09-724-676-68213	Sequence 68213, A

ALIGNMENTS

RESULT 1
US-09-724-676-61358
; Sequence 61358, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: CompuGen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 CompuGen
; CURRENT APPLICATION NUMBER: US/09/724.676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 61358
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-61358

Query Match 93.0%; Score 970.5; DB 5; Length 197;
Best Local Similarity 93.4%; Pred. No. 1.6e-89;
Matches 185; Conservative 7; Mismatches 5; Indels 1; Gaps 1;

Qy	1	MDRFDYSNFSLLIACVANDVSESETRAKFESLFRDYDKDTTFOYFKSKRVINFSN	60
Db	1	MHFRNFYSFSLIACVANSDFSESETRAKFESLFRDYDKDTTFOYFKSKRVINFSN	60
Qy	61	PLSAADARLRHLKTEFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPPVGVKQV	120
Db	61	PFSAADARLRHLKTEFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPPVGVKQV	120
Qy	121	EDATPVINYDLYLAISKLGPEKVELHAATDPTPSVVHVHVCESQDNEEEEMERMKRP	180
Db	121	EDATPVINYDLYLAISKLGPEKVELHAATDPTPSVVHVHVCESQDNEEEEMERMKRP	179
Qy	181	KPKLIQTRRPEYTPHLS	198
Db	180	KPKLIQTRRPEYTPHLS	197

RESULT 2
US-09-724-676-61359
; Sequence 61359, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: CompuGen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 CompuGen
; CURRENT APPLICATION NUMBER: US/09/724.676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2

RESULT 6
US-09-724-676A-61367
; Sequence 61367, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD


```

; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 61354
; LENGTH: 171
; TYPE: PR1
; ORGANISM: Homo sapiens
US-09-724-676A-61354

```

```

Query Match      81.3%; Score 847.5; DB 5; Length 171;
Best Local Similarity 95.9%; Pred. No. 2,6e-77;
Matches 162; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

```

```

Qy 30 AKFESLFRYYDKOTTFQYFKSFKRVIRINFSNPLSADARLRLHKTFFLGKEMKLYFAQTL 89
Db 4 AKFESLFRYYDKOTTFQYFKSFKRVIRINFSNPLSADARLRLHKTFFLGKEMKLYFAQTL 63
Qy 90 HIGSSHLAPPNPDKQFLISPPASPPVGMKQVEDATPVINYDLYLAISKLGPEKYEELHAA 149
Db 64 HIGSSHLAPPNPDKQFLISPPASPPVGMKQVEDATPVINYDLYLAISKLGPEKYEELHAA 123
Qy 150 TDPPTSVVVHVCSDOENEEBEEEMERMRKPKKIOTRRPEYTPIHLS 198
Db 124 TDTPTSVVVHVCSDOE-KEEEMERMRMRPKKIOTRRPEYTPIHLS 171

```

```

RESULT 11
US-09-724-676-61356
; Sequence 61356, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 61356
; LENGTH: 200
; TYPE: PR1
; ORGANISM: Homo sapiens
US-09-724-676-61356

```

```

Query Match      81.3%; Score 847.5; DB 5; Length 200;
Best Local Similarity 95.9%; Pred. No. 3,2e-77;
Matches 162; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

```

```

Qy 30 AKFESLFRYYDKOTTFQYFKSFKRVIRINFSNPLSADARLRLHKTFFLGKEMKLYFAQTL 89
Db 33 AKFESLFRYYDKOTTFQYFKSFKRVIRINFSNPLSADARLRLHKTFFLGKEMKLYFAQTL 92
Qy 90 HIGSSHLAPPNPDKQFLISPPASPPVGMKQVEDATPVINYDLYLAISKLGPEKYEELHAA 149
Db 93 HIGSSHLAPPNPDKQFLISPPASPPVGMKQVEDATPVINYDLYLAISKLGPEKYEELHAA 152
Qy 150 TDPPTSVVVHVCSDOENEEBEEEMERMRKPKKIOTRRPEYTPIHLS 198
Db 153 TDTPTSVVVHVCSDOE-KEEEMERMRMRPKKIOTRRPEYTPIHLS 200

```

```

RESULT 12
US-09-724-676A-61356
; Sequence 61356, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A

```

```

; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 61356
; LENGTH: 200
; TYPE: PR1
; ORGANISM: Homo sapiens
US-09-724-676A-61356

```

```

Query Match      81.3%; Score 847.5; DB 5; Length 200;
Best Local Similarity 95.9%; Pred. No. 3,2e-77;
Matches 162; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

```

```

Qy 30 AKFESLFRYYDKOTTFQYFKSFKRVIRINFSNPLSADARLRLHKTFFLGKEMKLYFAQTL 89
Db 33 AKFESLFRYYDKOTTFQYFKSFKRVIRINFSNPLSADARLRLHKTFFLGKEMKLYFAQTL 92
Qy 90 HIGSSHLAPPNPDKQFLISPPASPPVGMKQVEDATPVINYDLYLAISKLGPEKYEELHAA 149
Db 93 HIGSSHLAPPNPDKQFLISPPASPPVGMKQVEDATPVINYDLYLAISKLGPEKYEELHAA 152
Qy 150 TDPPTSVVVHVCSDOENEEBEEEMERMRKPKKIOTRRPEYTPIHLS 198
Db 153 TDTPTSVVVHVCSDOE-KEEEMERMRMRPKKIOTRRPEYTPIHLS 200

```

```

RESULT 13
US-09-724-676-61360
; Sequence 61360, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 61360
; LENGTH: 205
; TYPE: PR1
; ORGANISM: Homo sapiens
US-09-724-676-61360

```

```

Query Match      67.6%; Score 705; DB 5; Length 205;
Best Local Similarity 93.0%; Pred. No. 5,7e-63;
Matches 133; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

```

```

Qy 1 MDRFDSYNSFLIACVANDDVSESETRAKFESLFRYYDKOTTFQYFKSFKRVIRINFSN 60
Db 1 MAFRPNYSFSSLICVANSDFSESETRAKFESLFRYYDKOTTFQYFKSFKRVIRINFSN 60
Qy 61 PLSADARLRLHKTFFLGKEMKLYFAQTLHIGSSHLAPPNPDKQFLISPPASPPVGMKQV 120
Db 61 PLSADARLRLHKTFFLGKEMKLYFAQTLHIGSSHLAPPNPDKQFLISPPASPPVGMKQV 120
Qy 121 EDPATPVINYDLYLAISKLGPEK 143
Db 121 EDPATPVINYDLYLAISKLGPEK 143

```

```

RESULT 14
US-09-724-676A-61360
; Sequence 61360, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 61360
; LENGTH: 205

```


THIS PAGE BLANK (UBPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 14, 2002, 20:09:21 ; Search time 52.7555 Seconds
(without alignments)
2732.361 Million cell updates/sec

Title: US-09-782-953-7

Perfect score: 1043

Sequence: 1 MDRFSYNFSLIACVAND.....RPKPKIIQRRRDEPTPIHLS 198

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 422006 seqs, 364007625 residues

Total number of hits satisfying chosen parameters: 844012

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODELS=frame+ p2n.model -DBV=xlh
-Q=/cgn2_1/USPTO.spool/US09782953/runat_11122002_114431_17431/app_query.fasta_1.1173
-DB=Pending Patents NA New -QFWT=fastcap -SUFFIX=inn -MINMATCH=0 -1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFWT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09782953 @C@N 1 1 70 @runat_11122002_114431_17431 -NCPU=6 -ICPU=3
-NO XLPXY -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Pending Patents NA New:

1: /cgn2_6/prodata/1/pna/PCT_NEW_COMB.seq.*
2: /cgn2_6/prodata/1/pna/US06_NEW_COMB.seq.*
3: /cgn2_6/prodata/1/pna/US07_NEW_COMB.seq.*
4: /cgn2_6/prodata/1/pna/US08_NEW_COMB.seq.*
5: /cgn2_6/prodata/1/pna/US09_NEW_COMB.seq.*
6: /cgn2_6/prodata/1/pna/US10_NEW_COMB.seq.*
7: /cgn2_6/prodata/1/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	970.5	93.0	1931	5	US-09-724-676-12748 Sequence 12748, A
2	970.5	93.0	1931	5	US-09-724-676-12748 Sequence 12748, A
3	970.5	93.0	2050	5	US-09-724-676-12756 Sequence 12756, A
4	970.5	93.0	2050	5	US-09-724-676-12756 Sequence 12756, A
5	970.5	93.0	2297	5	US-09-724-676-12747 Sequence 12747, A
6	970.5	93.0	2297	5	US-09-724-676-12747 Sequence 12747, A
7	957.5	91.8	2348	6	US-10-240-965-60 Sequence 60, Appl
8	874	83.8	1893	5	US-09-724-676-12741 Sequence 12741, A
9	874	83.8	1893	5	US-09-724-676-12741 Sequence 12741, A
10	850.5	81.5	1876	5	US-09-724-676-12745 Sequence 12745, A
11	850.5	81.5	1876	5	US-09-724-676-12745 Sequence 12745, A

12	847.5	81.3	1839	5	US-09-724-676-12743 Sequence 12743, A
13	847.5	81.3	1839	5	US-09-724-676-12743 Sequence 12743, A
14	705	67.6	911	5	US-09-724-676-12749 Sequence 12749, A
15	705	67.6	911	5	US-09-724-676-12749 Sequence 12749, A
16	702	67.3	1094	5	US-09-724-676-12750 Sequence 12750, A
17	702	67.3	1094	5	US-09-724-676-12750 Sequence 12750, A
18	702	67.3	1213	5	US-09-724-676-12740 Sequence 12740, A
19	702	67.3	1213	5	US-09-724-676-12740 Sequence 12740, A
20	623	59.7	3159	6	US-10-290-438-1 Sequence 1, Appl
21	605.5	58.1	1056	5	US-09-724-676-12742 Sequence 12742, A
22	605.5	58.1	1056	5	US-09-724-676-12742 Sequence 12742, A
23	597.5	57.3	2411	5	US-09-724-676-15329 Sequence 15329, A
24	597.5	57.3	2411	5	US-09-724-676-15329 Sequence 15329, A
25	582	55.8	1039	5	US-09-724-676-12746 Sequence 12746, A
26	582	55.8	1039	5	US-09-724-676-12746 Sequence 12746, A
27	579	55.5	1002	5	US-09-724-676-12744 Sequence 12744, A
28	579	55.5	1002	5	US-09-724-676-12744 Sequence 12744, A
29	537	51.5	442	6	US-09-513-999C-1772 Sequence 1772, Ap
30	536.5	51.4	615	6	US-10-290-438-8 Sequence 8, Appl
31	528.5	50.7	1021	6	US-10-290-438-3 Sequence 3, Appl
32	415.5	39.8	1577	5	US-09-724-676-12755 Sequence 12755, A
33	415.5	39.8	1577	5	US-09-724-676-12755 Sequence 12755, A
34	415.5	39.8	1943	5	US-09-724-676-12752 Sequence 12752, A
35	415.5	39.8	1943	5	US-09-724-676-12752 Sequence 12752, A
36	285.5	27.4	446	6	US-10-203-138A-2101 Sequence 2101, Ap
37	266.5	25.6	486	6	US-10-203-138A-827 Sequence 827, App
38	198.5	19.0	531	6	US-10-290-438-6 Sequence 6, Appl
39	187.5	18.0	123	6	US-10-203-138A-5980 Sequence 5980, Ap
40	170	16.3	465	5	US-09-620-607B-1320 Sequence 1320, Ap
41	150	14.4	557	5	US-09-724-676-12753 Sequence 12753, A
42	150	14.4	557	5	US-09-724-676-12753 Sequence 12753, A
43	147	14.1	740	5	US-09-724-676-12754 Sequence 12754, A
44	147	14.1	740	5	US-09-724-676-12754 Sequence 12754, A
45	140	13.4	85	6	US-10-203-138A-7216 Sequence 7216, Ap

ALIGNMENTS

RESULT 1
US-09-724-676-12748
; Sequence 12748, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: CompuGen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 CompuGen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12748
; LENGTH: 1931
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (11)..(11)
; OTHER INFORMATION: n is a,c,g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (18)..(18)
; OTHER INFORMATION: n is a,c,g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (23)..(23)
; OTHER INFORMATION: n is a,c,g, or t
US-09-724-676-12748
Alignment Scores:
Pred. No.: 4.31e-100 Length: 1931
Score: 970.50 Matches: 185
Percent Similarity: 96.97% Conservative: 7
Best Local Similarity: 93.43% Mismatches: 5
Query Match: 93.05% Indels: 1

DB: 5 Gaps: 1

US-09-782-953-7 (1-198) x US-09-724-676-12748 (1-1931)

QY 1 MetAapPheArpAapPheSerTyRAsnPheSerSerLeuIleAlaCyValAlaAsnAap 20

DB 63 ATGCATTTTGAAGAACTTAACTTACATTAAGCTCCCTGATTCCTGCTGGCAACAGT 122

QY 21 AapValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThrTyRAsp 40

DB 123 GATATCTTCAGCGAAAGTAAACAGGGCCAAATTGATCCCTCTTAGAGCGATAGC 182

QY 41 LysAapThrThrPheGlnTyRPhelySerPheLysValArgIleAsnPheSerAsn 60

DB 183 AAGGACATCACCTTTCAGATTTTAAAGACTTCAAAAGAGTCAGATTAACCTTCAGCAAC 242

QY 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80

DB 243 CCTTCTCCGAGCAGATGCCAGGCTCCAGCTGCATAGACTGAGTTCTGGAAAGGA 302

QY 81 MetLysLeuTyRPhelArgInThrLeuHisIleGlySerSerHisLeuAlaProProAsn 100

DB 303 ATGAACTTATATTTCCTCAGACCTTACATAGAAAGCTCACACCTGGCTCCGCAAAAT 362

QY 101 ProAapLysGlnPheLeuIleSerProProAlaSerProProValGlyTyRplyGlnVal 120

DB 363 CCAGACAAAGCAGTTTGTGATCTCCCTCCGCTCCGCGCAGTGGAGTGAACAAAGTG 422

QY 121 GluAapAlaThrProValIleAsnTyRAspLeuLeuTyRAlaIleSerLysLeuGlyPro 140

DB 423 GAAGAGCCAGCCCACTCAATAACTATGATCTCTTAATAGCCATCTCCAAAGCTGGGGCCA 482

QY 141 GlyLysLysTyRGlulLeuHisAlaAlaThrAspProThrProSerValValValHisVal 160

DB 483 GGGGAAAGTATGAATTGACGCGAGCTGACACCACTCCAGCGTGGTGGCATGTA 542

QY 161 CygGluSerAspGlnGluAsnGluGluGluGluGluGluGluGluGluGluGluGluGlu 180

DB 543 TGTGAGAGTGATCAAGAG--AAGGAGGAAGAAAGGAAATGGAAATGAGAGACCT 599

QY 181 LysProLysIleIleGlnThrArgArgProGluTyRThrProIleHisLeuSer 198

DB 600 AAGCCAAAATTTTCCAGACCAAGAGGCCGAGTACCGCCGATCCACTCAGC 653

RESULT 2

US-09-724-676A-12748

Sequence 12748, Application US/09724676A

GENERAL INFORMATION:

APPLICANT: Compugen LTD

TITLE OF INVENTION: Variants of alternative splicing

FILE REFERENCE: 129181.4 Compugen

CURRENT APPLICATION NUMBER: US/09/724,676A

CURRENT FILING DATE: 2000-11-28

NUMBER OF SEQ ID NOS: 97222

SOFTWARE: PatentIn version 3.2

SEQ ID NO 12748

LENGTH: 1931

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc_feature

LOCATION: (18)..(18)

OTHER INFORMATION: n is a,c,g, or t

FEATURE:

NAME/KEY: misc_feature

LOCATION: (23)..(23)

OTHER INFORMATION: n is a,c,g, or t

US-09-724-676A-12748

Alignment Scores:

Pred. No.: 4,31e-100 Length: 1931

Score: 970.50 Matches: 185

Percent Similarity: 96.97% Conservative: 7

Best Local Similarity: 93.43% Mismatches: 5

Query Match: 93.05% Indels: 1

DB: 5 Gaps: 1

US-09-782-953-7 (1-198) x US-09-724-676A-12748 (1-1931)

QY 1 MetAapPheArpAapPheSerTyRAsnPheSerSerLeuIleAlaCyValAlaAsnAap 20

DB 63 ATGCATTTTGAAGAACTTAACTTACATTAAGCTCCCTGATTCCTGCTGGCAACAGT 122

QY 21 AapValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThrTyRAsp 40

DB 123 GATATCTTCAGCGAAAGTAAACAGGGCCAAATTGATCCCTCTTAGAGCGATAGC 182

QY 41 LysAapThrThrPheGlnTyRPhelySerPheLysValArgIleAsnPheSerAsn 60

DB 183 AAGGACATCACCTTTCAGATTTTAAAGACTTCAAAAGAGTCAGATTAACCTTCAGCAAC 242

QY 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80

DB 243 CCTTCTCCGAGCAGATGCCAGGCTCCAGCTGCATAGACTGAGTTCTGGAAAGGA 302

QY 81 MetLysLeuTyRPhelArgInThrLeuHisIleGlySerSerHisLeuAlaProProAsn 100

DB 303 ATGAACTTATATTTCCTCAGACCTTACATAGAAAGCTCACACCTGGCTCCGCAAAAT 362

QY 101 ProAapLysGlnPheLeuIleSerProProAlaSerProProValGlyTyRplyGlnVal 120

DB 363 CCAGACAAAGCAGTTTGTGATCTCCCTCCGCTCCGCGCAGTGGAGTGAACAAAGTG 422

QY 121 GluAapAlaThrProValIleAsnTyRAspLeuLeuTyRAlaIleSerLysLeuGlyPro 140

DB 423 GAAGAGCCAGCCCACTCAATAACTATGATCTCTTAATAGCCATCTCCAAAGCTGGGGCCA 482

QY 141 GlyLysLysTyRGlulLeuHisAlaAlaThrAspProThrProSerValValValHisVal 160

DB 483 GGGGAAAGTATGAATTGACGCGAGCTGACACCACTCCAGCGTGGTGGCATGTA 542

QY 161 CygGluSerAspGlnGluAsnGluGluGluGluGluGluGluGluGluGluGluGluGlu 180

DB 543 TGTGAGAGTGATCAAGAG--AAGGAGGAAGAAAGGAAATGGAAATGAGATGAGAGACCT 599

QY 181 LysProLysIleIleGlnThrArgArgProGluTyRThrProIleHisLeuSer 198

DB 600 AAGCCAAAATTTTCCAGACCAAGAGGCCGAGTACCGCCGATCCACTCAGC 653

RESULT 3

US-09-724-676-12756

Sequence 12756, Application US/09724676

GENERAL INFORMATION:

APPLICANT: Compugen LTD

TITLE OF INVENTION: Variants of alternative splicing

FILE REFERENCE: 129181.4 Compugen

CURRENT APPLICATION NUMBER: US/09/724,676

CURRENT FILING DATE: 2000-11-28

NUMBER OF SEQ ID NOS: 97222

SOFTWARE: PatentIn version 3.2

SEQ ID NO 12756

LENGTH: 2050

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc_feature

LOCATION: (7)..(7)

OTHER INFORMATION: n is a,c,g, or t

FEATURE:

NAME/KEY: misc_feature

LOCATION: (47)..(47)

OTHER INFORMATION: n is a,c,g, or t

US-09-724-676-12756

Alignment Scores:

Pred. No.:	4.68e-100	Length:	2050
Score:	970.50	Matches:	185
Percent Similarity:	96.97%	Conservative:	7
Best Local Similarity:	93.43%	Mismatches:	5
Query Match:	93.05%	Indels:	1
DB:	5	Gaps:	1

US-09-782-953-7 (1-198) x US-09-724-676-12756 (1-2050)

Qy	1	MetAspPheArgAspPheSerTyrAsnPheSerSerLeuIleAlaCysValAlaAsnAsp	20
Db	182	ATGCATTTTAAAGAACTTAACTACAGCTTTTAGCTCCCTGATTCCTGTGGCAACAGT	241
Qy	21	AspValPheSerGluSerGluThrAtqAlaValPheGluSerLeuPheArgThrTyrAsp	40
Db	242	GATATCTTCAGCNAAGTGAACCCAGGGCCAAATTTGAGTCCCTCTTTAGGAGCGTATGAC	301
Qy	41	LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAenPheSerAsn	60
Db	302	AAGGACATCACCTTTTCAGTATTTTAAGAGCTTCAACAGCAGTCAGATAAATTCACGCAAC	361
Qy	61	ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu	80
Db	362	CCCTTCTCCGAGCAGATGCCAGCTCCAGCTGCATGAAGACTCAGTTTCTGGGAAGGAA	421
Qy	81	MetLysLeuTyrPheAlaGlnThrLeuHisGlyGlySerSerHisLeuAlaProProAsn	100
Db	422	ATGAAGTTATATTTTGTCTCAGACCTTACACATAGGAAGCTCACACTGGCTCCGCCAAAT	481
Qy	101	ProAspLysGlnPheLeuIleSerProProAlaSerProProValGlyTyrLysGlnVal	120
Db	482	CCAGACAAGCAGTTTCTGATCTCCCTCCGCTCTCCGCGAGTGGATGGAAACAAGTG	541
Qy	121	GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro	140
Db	542	GAAGATCGACCCCGAGTCATAAACTATGATCTCTTATATGCCATCTCCAAGCTGGGGCCA	601
Qy	141	GlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal	160
Db	602	GGGGAAGAAGTATGATTTGCACGACGCGACTGCACACACTCCCAGCGTGGTGGTCCATGTA	661
Qy	161	CysGluSerAspGlnGluAsnGluGluGluGluMetGluArgMetLysArgPro	180
Db	662	TGTGAGAGTGCATCAAGAG---AAGGAGGAGAGAGAGAAATGAAAGAAATGAGCAGACCT	718
Qy	181	LysProLysIlelleGlnThrArgArgProGluTyrThrProIleHisLeuSer	198
Db	719	AAGCCAAAATTTATCCAGACAGAGAGCGGAGTACACGCCGATCCACCTCAGC	772

RESULT 4

US-09-724-676A-12756
; Sequence 12756, Application US/09724676A

```

;
; GENERAL INFORMATION:
;
; APPLICANT: CompuGen LTD
;
; TITLE OF INVENTION: Variants of alternative splicing
;
; FILE REFERENCE: 123181.4 CompuGen
;
; CURRENT APPLICATION NUMBER: US/09/724.676A
;
; CURRENT FILING DATE: 2000-11-28
;
; NUMBER OF SEQ ID NOS: 97222
;
; SOFTWARE: PatentIn version 3.1.2
;

```

; LOCATION: (47)..(47)
 ; OTHER INFORMATION: n is a,c,g, or t
 US-09-724-676A-12756

Alignment Scores:	
Pred. No.:	4.68e-100
Score:	970.50
Percent Similarity:	96.97%
Best Local Similarity:	93.43%
Query Match:	93.05%
DB:	5
Length:	2050
Matches:	185
Conservative:	7
Mismatches:	5
Indels:	1
Gaps:	1

US-09-782-953-7 (1-198) x US-09-724-676A-12756 (1-2050)

Qy	1	MetAspPheArgAspPheSerTyrAsnPheSerSerLeuIleAlaCysValAlaAsnAsp	20
Db	182	ATGCATTTTAAAGAACTTAACTACAGCTTTTAGCTCCCTGATTCCTGTTGTGGCAACAGT	241
Qy	21	AspValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThrTyrAsp	40
Db	242	GATATCTTTCAGGAAAGTGAACACAGGCCAAATTTGAGTCCCTCTTTAGGACGTATGAC	301
Qy	41	LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn	60
Db	302	AAGGACATCACCTTCTCAGTATTTTAAGAGCTTCAACAGAGTCAGATNAACTTCAGCAAC	361
Qy	61	ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu	80
Db	362	CCCTTCTCCGAGCAGATGCAGGCTCCAGCTGCATAAGACTGAGTTTCTGGGAAGGAA	421
Qy	81	MetLysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProPheAsn	100
Db	422	ATGAAGTTATATTGTCTCAGACCTTACACATAGGAAGCTCACACTGGCTCGCCAAAT	481
Qy	101	ProAspLysGlnPheLeuIleSerProProAlaSerProProValGlyTrpLysGlnVal	120
Db	482	CCAGACAAGCAGTTCTCGATCTCCCTCCGCTCTCCGCCAGTGGGATGGAACAAGTG	541
Qy	121	GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro	140
Db	542	GAAGATGCGAGCCCGAGTCATAAACTATGATCTCTTATATGCCATCTCCAAGCTGGGGCCA	601
Qy	141	GlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal	160
Db	602	GGGGAAAGTATGAATTGCACGGCGGACTTGACCCACTCCCAAGCTGGTGGTCCATGTA	661
Qy	161	CysGluSerAspGlnGluAsnGluGluGluGluMetGluArgMetLysArgPro	180
Db	662	TGTGAGATGATCAAGAG --- AAGGAGGAAGAGAGAGGAAATGGAAAGATGAGGAGACCT	718
Qy	181	LysProLysIleIleGlnThrArgArgProGluTyrThrProIleHisLeuSer	198
Db	719	AAGCCAAAATTTATCAGACCGAGGAGCGGAGTACACGCCGATCCACCTCAGC	772

RESULT, T 5

US-09-724-676-12747
: Sequence 12747. Application US/09724676

```

; GENERAL INFORMATION:
; APPLICANT: CompuGen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 CompuGen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2

```

```
;
;
FEATURE:
NAME/KEY: misc feature
```

```

? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (18)_.(18)
? OTHER_INFORMATION: n is a,c,g, or t
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (23)_.(23)
? OTHER_INFORMATION: n is a,c,g, or t
US-09-724-676-12747

```

Alignment Scores:	
Pred. No.:	5.47e-100
Score:	970.50
Percent Similarity:	96.97%
Best Local Similarity:	93.43%
Query Match:	93.05%
DB:	5
Length:	2297
Matches:	185
Conservative:	7
Mismatches:	5
Indels:	1
Gaps:	1

US-09-782-953-7 (1-198) X US-09-724-676-12747 (1-22977)

Oy 1 MetAspPheArgAspPheSerTyrAsnProSerLeuIleAlaValAlaAsnAbp 20
|||:::||:::||
Db 63 ATGCATTTAGAACTTAACACGAGTTTAGCCTCCGTGAATTGGCAACAGT 122

Qy 21 AspvAlPheesrGIuseGIuThrArGaIaLyBPhieGIuserLeuBheArGThrTyAsp 40
||::|||
Db 123 GATATCTTCAGCGAAAGTGAACCCAGGCGCAATTGAGTCCCTCTTAGGACGTATGAC 182

Qy 41 LysAspThrIrrPheGIntYrPhelySsrPhelybArgValArgIleasnHesSerAsn 60
Db 183 AAGGACATCACTTCAGTATTTAAGAGCTCAACGAGTCAGAAATAACTTCAGCAAC 242

Oy 61 ProLeuSerAlaAlaAspAlaArgLeuAArgLeuHtIbVetThrGluPhenLeuGIylYpsGIu 80
DB 243 CCGCTTCCGGACAGATGCCAGGCTCAGCTGCATAAGACTGAGTTTCGTGGAAAGCAA 302

Qy 81 MetLysLeuYrPheAlaGlnThrLeuHisIecLysSerSerHisLeuAlaProProAsn 100
 Db 303 ATGAAGTATATTTTGGTCAGACCTTACACATAGGAAGCTCACCCTGGCTGGCCAAAT 362

Qy	101	ProApblybGInPheLeuIleSerProProAlaSerProProValGlyTrrpLySGlnVal	120
Db	363	CCAGACAAGCAGTTTCGATCTCCCTCCCGCCCTCCGCGCAGTGGGATGGAACAAGTG	422

QY 121 GUAAGAATATTTTProval11LeaenttyrAspLeuLeuTyrAla11LeuSerLybLeuGlyPro 140
DB 423 GAAGATGGCAACCCAGTCATAACTATGATCTCTTATATGTCATCTCCAAAGCTGGGCCA 483

QY 141 GYGLYSLYSTRGULEWHISALAAAThrAspProThrProSerValValValHisVal 166
DB 483 GGGGAAAGTATGAATTGCACGACGCACTGCCAGCGTGTGTCATGTA 542

Qy 161 CygBcUserarPcIngluaNglucInglucInglucInglucmetGluArgmetLysArgPro 180
| | | | | : | | | | | : | | | | |
Db 543 TGTGAGTGTATCAAGAG--AAGAGGACAGACGAATGGAAAGATGACGAGACTT 59

Qy 181 LysProLysaileilecInThrArgArgProGluYrThrProIleHisLeuSer 198
Db 600 AAGCAAAATATTCAGACACGAGGCGGAGTACACGCCGATCCACTCAGC 653

RESULT 6
US-09-724-676A-12747
: Sequence 12747, Application US/09724676A

```
; APPLICANT: Comugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Comugen
```

```

; CURRENT FILING DATE: 2000-11-28
;
; NUMBER OF SEQ ID NOS: 97222
;
; SOFTWARE: PatentIn version 3.2

```

; LENGTH: 2297

```

? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (11)..(11)
? OTHER INFORMATION: n is a,c,g, or t
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (18)..(18)
? OTHER INFORMATION: n is a,c,g, or t
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (23)..(23)
? OTHER INFORMATION: n is a,c,g, or t
US-09-124-676A-12747

```

Alignment Scores:	
Pred. No.:	5, 47e-100
Score:	970.50
Percent Similarity:	96.97%
Best Local Similarity:	93.44%
Query Match:	93.05%
DB:	5
Length:	2297
Matches:	185
Conservative:	7
Mismatches:	5
Indels:	1
Gaps:	1

US-09-782-953-7 (1-198) X US-09-724-676A-12747 (1-22977)

Oy 1 MetasphearigaspheSerTyrrasnPieSerSerLeuIealCyvalAlaasnAsp 20
 ||| ||||| ::::| ::::| ::::| ::::| ::::| ::::| ::::| ::::| ::::| ::::|
 Db 63 ATGCATTTTAGAAACCTTAACCTACAGCTTTAGCTCCCTGATTGGCCTGTGGCAACAGT 122

```

Oy 21 AspValPheSerGluSerGluThrTrpArgAlaLysPheGluSerLeuPheArgThrTyrAsp 400
    |||::|||
Db 123 GATATCTTCAGCGAAAGTGAACCAAGCGCAATTTGAGTCCCTTTAGGACGATGAC 180

```

Qy 41 LysaspThrThrPhnegIntYrPhelySserPhelysArgValArgIleasnPheseAsn 600
Db 183 AAGGACATCACCTTTCAGTATTTTAAGAGCTTCAACGAGTCAGATAAAGCTTCAGCAAC 240

Oy 61 ProLeuSerAlaAlaSpAlaArgLeuArgLeuHsLysThrGluPheLeuGlyLysGlu 800
 |||
 Db 243 CCCTTCTCCGAGCAGATGCCAGGCTCCAGCTGCATTAAGACTGAGTTCTGGGAAAGAA 300

Dy 61 MetLysLeuTyrrPhealaglInthrLeuhslleglySerSerHisleualaproProAsn 10
Db 303 ATGAAGTATATTGCTCAGACCTTACACATAGAAGCTCACACCTGGCTCCGCCAAT 366

Qy	101	ProAspLysGlnPheLeuIleSerProProAlaSerProProValGlyTTrpLysGlnVal	122
Db	363	CCAGACAAGACGTTTCGATCTCCCTCCGCGCTCTCCGCGAGTGGATGGAAACAAGTG	422

Qy 121 GLNAPPAIATHRPVVALI^{LE}ANTYRAAPLEULENTYRALAI^{ES}ETRYSLBLEUGI^YPRO 14
Db 423 GAAGATGGCA^{CC}CACTCATAACTATGATCTCTATATGCATCTCCAAGCTGGGGCCA 48

Oy 141 GlyGlyLeuYrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 16
|||
Db 483 GGGAAAGTAGATTGCACCGCAGCATGCACCACCTCCGCGTGTGTTCATGTA 54

Qy 161 CygglUserarpglngluAanglucnglucnglumeccluarqmetLysaaqPro 18
| | | | | : | | | | | : | | | | |
Db 543 TGTGAGTGATCAAGAG--AAAGAGAAGAAAGAGAAATGGAAAAGATTGAGAGACCT 59

QY 181 LysProlystleileglnThrArgArgProglutylThrProleHisbleuSer 198
 Db 600 AAGCAAAATATTCAGACAGAGAGCGAGTACAGCCGATCCACCTCAGC 653

RESULT 7
US-10-240-965-60
; Sequence 60, Application US/10240965

```

; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: SHIFFMAN, DOV
; APPLICANT: SOMOGYI, Roland
;

```

APPLICANT: SOMOGYI, Roland

; APPLICANT: LAWN, Richard M.
 ; APPLICANT: SEILHAMER, Jeffrey J.
 ; APPLICANT: PORTER, Gordon J.
 ; APPLICANT: MIKITA, Thomas
 ; APPLICANT: TAI, Julie
 ; TITLE OF INVENTION: GENES EXPRESSED IN FOMM CELL DIFFERENTIATION
 ; CURRENT APPLICATION NUMBER: US/10/240,965
 ; CURRENT FILING DATE: 2002-10-04
 ; PRIOR APPLICATION NUMBER: 60/195,106
 ; PRIOR FILING DATE: 2000-04-05
 ; NUMBER OF SEQ ID NOS: 276
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 60
 ; LENGTH: 2348
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; OTHER INFORMATION: incyte ID No: 042176.5
 US-10-240-965-60

Alignment Scores:
 Pred. No.: 1,65e-98 Length: 2348
 Score: 957.50 Matches: 185
 Percent Similarity: 96.48% Conservative: 7
 Best Local Similarity: 92.96% Mismatches: 5
 Query Match: 91.80% Indels: 2
 DB: 6 Gaps: 1

US-09-782-953-7 (1-198) x US-10-240-965-60 (1-2348)

QY 1 MetAspPheArgAspPheSerTyrAsnPheSerSerLeuIleAlaCysValAlaAsnAsp 20
 DB 151 ATGCATTTTAGAACTTTAACTACAGTTTAGCTCCCTGATTCCTGTGTGGCAACAGT 210
 QY 21 AspValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40
 DB 211 GATATCTTCAGCAAGTGAACCCAGGGCCAAATTTGAGTCCCTCTTTAGGAGGTATGAC 270
 QY 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60
 DB 271 AAGGACATCACCCTTCAGTATTTAAGAGCTTCAACAGCTCAGATTAACCTTCAGCAAC 330
 QY 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
 DB 331 CCCTTCTCCGACGAGATGCCAGCTCCAGCTGCATTAAGACTCAGTTTCTGGGAAAGGAA 390
 QY 81 MetLysLeuTyrPheAlaGlnThrLeuHisLysGlySerSerHisLeuAlaProProAs 100
 DB 391 ATGAGTTATATTTGCTCAGACCTTACACATAGGAAGCTCACACCTGGGCTCCGCCAA 450
 QY 100 nProAspLysGlnPheLeuIleSerProProAlaSerProProValGlyTyrLysGlnVa 120
 DB 451 TCCAGACAAGCAGTTTCTGATCTCCCTCCCGCTCTCCGCCAGTGGGATGGAAACAGT 510
 QY 120 lGluAspAlaThrProValIleAsnTyrAspLeuTyrAlaIleSerLysLeuGlyPr 140
 DB 511 GGAAGATGCGACCCAGTATTAACATATGATCTTATATGCCATCTCAAGCTGGGGCC 570
 QY 140 oGlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVa 160
 DB 571 AGGGGAAAGTATGAATTGACCCAGGACTGACACCTCCCGAGGTGGTGGTCCATGT 630
 QY 160 lCysGluSerAspGlnGluAsnGluGluGluGluMetGluArgMetLysArgPr 180
 DB 631 ATGTGAGTGTATCAAGAG---AAGGAGGAAGAGGAAATGGAAAGATGAGGAGACC 687
 QY 180 oLysProLysIleLeuThrArgArgProGluTyrThrProIleHisLeuSer 198
 DB 688 TAAGCCAAAATTTATCCAGACCGAGGCGGAGTAGTACGCCGCTCAGCTCAGC 742

RESULT 8

US-09-724-676-12741
 ; Sequence 12741, Application US/09724676
 ; GENERAL INFORMATION:
 ; APPLICANT: Compugen LTD
 ; TITLE OF INVENTION: Variants of alternative splicing
 ; FILE REFERENCE: 129181.4 Compugen
 ; CURRENT APPLICATION NUMBER: US/09/724,676
 ; CURRENT FILING DATE: 2000-11-28
 ; NUMBER OF SEQ ID NOS: 97222
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 12741
 ; LENGTH: 1893
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-724-676-12741

Alignment Scores:
 Pred. No.: 3.21e-89 Length: 1893
 Score: 874.00 Matches: 171
 Percent Similarity: 89.90% Conservative: 7
 Best Local Similarity: 86.36% Mismatches: 16
 Query Match: 83.80% Indels: 4
 DB: 5 Gaps: 2

US-09-782-953-7 (1-198) x US-09-724-676-12741 (1-1893)

QY 1 MetAspPheArgAspPheSerTyrAsnPheSerSerLeuIleAlaCysValAlaAsnAsp 20
 DB 34 GTGGACCTCGAGGACCTGCC-----AGCGCCACCATCGCTGTCCACCTGGACCG 84
 QY 21 AspValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40
 DB 85 CCGGTGTTGTCGACGGCTGTGCGGCCAAATTTAGTCCCTCTTTAGGACGTATGAC 144
 QY 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60
 DB 145 AAGGACATCACCCTTTTCAAGAGCTTCAACAGCTCAGATTAACCTTCAGCAAC 204
 QY 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
 DB 205 CCCTTCTCCGACGAGATGCCAGCTCCAGCTGCATTAAGACTGAGTTTCTGGGAAAGGAA 264
 QY 81 MetLysLeuTyrPheAlaGlnThrLeuHisLysGlySerSerHisLeuAlaProProAsn 100
 DB 265 ATGAAGTTATATTTGCTCAGACCTTACACATAGGAAGCTCACACCTGGCTCCGCCAA 324
 QY 101 ProAspLysGlnPheLeuIleSerProProAlaSerProProValGlyTyrLysGlnVal 120
 DB 325 CCAGACAAGCAGTTTCTGATCTCCCTCCCGCTCTCCGCCAGTGGGATGGAAACAAGTG 384
 QY 121 GluAspAlaThrProValIleAsnTyrAspLeuTyrAlaIleSerLysLeuGlyPro 140
 DB 395 GAAGATGCGACCCAGTATTAACATATGATCTTATATGCCATCTCAAGCTGGGGCCA 444
 QY 141 GlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160
 DB 445 GGGGAAAGTATGAATTTCACGACGAGCTCACACCTCCCGAGGTGGTGGTCCATGTA 504
 QY 161 CysGluSerAspGlnGluAsnGluGluGluGluMetGluArgMetLysArgPro 180
 DB 505 TGTGAGAGTGTATCAAGAG---AAGGAGGAAGAGGAAATGGAAAGATGAGGAGACCT 561
 QY 181 LysProLysIleLeuGlnThrArgArgProGluTyrThrProIleHisLeuSer 198
 DB 562 AAGCCAAAATTTTCCAGACCGAGGCGGAGTAGTACGCCGCTCAGCTCAGC 615

RESULT 9
 US-09-724-676A-12741
 ; Sequence 12741, Application US/09724676A
 ; GENERAL INFORMATION:
 ; APPLICANT: Compugen LTD
 ; TITLE OF INVENTION: Variants of alternative splicing
 ; FILE REFERENCE: 129181.4 Compugen

LENGUIN: 10/0

Alignment Scores:	1.42e-86	Length:	1876
Pred. No.:			

Alignment scores:	
Pred. No.:	1.42e-86
Length:	1876

Score: 850.50 Matches: 165
 Percent Similarity: 93.48% Conservatives: 7
 Best Local Similarity: 89.67% Mismatches: 10
 Query Match: 81.54% Indels: 2
 DB: 5 Gaps: 1

US-09-782-953-7 (1-198) x US-09-724-676A-12745 (1-1876)

Qy 16 CysValAlaAsnAspValPheSerGluSerGlu-ThrArgAlaLysPheGluSerLe 35
 Db 52 TGCATTCCCAATCGCACTATTGGAAGAAAGCAATGTTGGGCCAAATTTGAGTCCCT 111
 Qy 35 upheArgThrTyraAspLysAspThrPheGlnTyPheLysSerPheLysArgValAr 55
 Db 112 CTTTAGACGATGACAGGACATCACCTTTCAGTATTTAAAGAGCTTCAAGAGTCAG 171
 Qy 55 gileAsnPheserAsnProLysSerAlaAlaAspAlaArgLeuHisLysThrGl 75
 Db 172 AATAAACTTCAGCAACCCCTTCTCCGAGCAGATGCCAGGCTCCAGCTGCATAAGACTGA 231
 Qy 75 upheLeuGlyLysGluMetLysLeuTyPheAlaGlnThrLeuHisLysGlySerSerHi 95
 Db 232 GTTTCCTGGGAAGAAATGAAGTTATATTTTGCTCAGACCTTACACATAGGAAGCTCACA 291
 Qy 95 sLeuAlaProAsnProAspLysGlnPheLeuLleSerProAlaSerProProva 115
 Db 292 CTGGCTCCGCCAATCCAGACAGCAAGTTCTGATCTCCCTCCCGCCCTCTCCGCGAGT 351
 Qy 115 lGlyTrpLysGlnValGluAspAlaThrProValLleAsnTyraAspLeuTyraLail 135
 Db 352 GGGATGGAACAAGTGAAGATGCGACCCAGTCATATAAATATGATCTCTTATATGCCAT 411
 Qy 135 eSerLysLeuGlyProGlyLysTyPheGluLeuHisAlaAlaThrAspProThrProse 155
 Db 412 CTCACAGCTGGGCCAGGGGAAAGATGAATTCGACGCGAGCTGACACCACTCCCGAG 471
 Qy 155 rValValHisValCysGluSerAspGlnGluAenGluGluGluGluMetGl 175
 Db 472 CTGGTGTCTCATGTATGTCAGAGTGATCAAGAG---AAGGAGAAAGAGGAAATGGA 528
 Qy 175 uArgMetLysArgProLysProLysLleGlnThrArgArgProGluTyPheThrProil 195
 Db 529 AAGAAATGAGGAGACCTAAGCAAAATATATCCAGCAGGAGCGCGAGTACAGCGCGAT 588
 Qy 195 eHisLeuSer 198
 Db 589 CCACCTCAGC 598

RESULT 12
 US-09-724-676-12743
 ; Sequence 12743, Application US/09724676
 ; GENERAL INFORMATION:
 ; APPLICANT: Compugen LTD
 ; TITLE OF INVENTION: Variants of alternative splicing
 ; FILE REFERENCE: 129181.4 Compugen
 ; CURRENT APPLICATION NUMBER: US/09/724,676
 ; CURRENT FILING DATE: 2000-11-28
 ; NUMBER OF SEQ ID NOS: 97222
 ; SOFTWARE: Patent in version 3.2
 ; SEQ ID NO 12743
 ; LENGTH: 1839
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-724-676-12743

Alignment Scores:
 Pred. No.: 3e-86 Length: 1839
 Score: 847.50 Matches: 162
 Percent Similarity: 97.63% Conservatives: 3
 Best Local Similarity: 95.86% Mismatches: 3
 Query Match: 81.26% Indels: 1
 DB: 5 Gaps: 1

US-09-782-953-7 (1-198) x US-09-724-676-12743 (1-1839)
 Qy 30 AlaLysPheGluSerLeuPheArgThrTyraAspLysAspThrPheGlnTyPheLys 49
 Db 58 GCCAAATTTGAGTCCCTCTTTAGGACGTATGACAGGACATCACCTTTTCAGTATTTTAAAG 117
 Qy 50 SerPheLysArgValArgLleAsnPheserAsnProLysSerAlaAlaAspAlaArgLeu 69
 Db 118 AGCTTCAACAGGATCAGATAAACTTCAGCAACCCCTTCTCCGAGCAGATGCCAGGCTC 177
 Qy 70 ArgLeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyPheAlaGlnThrLeu 89
 Db 178 CAGCTGCATAAGACTGAGTTCTTGGGAAAGAAATGAAGTTATATTTGCTCAGACCTTA 237
 Qy 90 HisLleGlySerSerHisLeuAlaProProAsnProAspLysGlnPheLeuLleSerPro 109
 Db 238 CACATAGGAGGCTCACCTGCTCGCCAAATCCAGACAGCAGTTTCTGATCTCCCT 297
 Qy 110 ProAlaSerProProValGlyTrpLysGlnValGluAspAlaThrProValLleAsnTyP 129
 Db 298 CCGCGCTCTCCGCCAGTGGGATGGAACAAGTGAAGATGCGACCCAGTCAATAACTAT 357
 Qy 130 AspLeuLeuTyPheAlaLleSerLysLeuGlyProGlyLysTyPheGluLeuHisAlaAla 149
 Db 358 GATCTCTTATATGCCATCTCCAAGCTGGGCCAGGGGAAAGATGATGAATTCACGACGCG 417
 Qy 150 ThrAspProThrProSerValValHisValCysGluSerAspGlnGluAenGluGlu 169
 Db 418 ACTGACACCACTCCAGCGTGGTGTCTCATGTATGTGAGATGATCAAGAG---AAGGAG 474
 Qy 170 GluGluGluGluMetGluArgMetLysArgProLysProLysLleGlnThrArgArg 189
 Db 475 GAAGAAGAGGAAATGGAAGAAATGAGGAGACCTAAGCAAAATATATCCAGACCAAGGAGG 534
 Qy 190 ProGluTyPheThrProLleHisLeuSer 198
 Db 535 CCGGAGTACACCGCATCCACCTCAGC 561

RESULT 13
 US-09-724-676A-12743
 ; Sequence 12743, Application US/09724676A
 ; GENERAL INFORMATION:
 ; APPLICANT: Compugen LTD
 ; TITLE OF INVENTION: Variants of alternative splicing
 ; FILE REFERENCE: 129181.4 Compugen
 ; CURRENT APPLICATION NUMBER: US/09/724,676A
 ; CURRENT FILING DATE: 2000-11-28
 ; NUMBER OF SEQ ID NOS: 97222
 ; SOFTWARE: Patent in version 3.2
 ; SEQ ID NO 12743
 ; LENGTH: 1839
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-724-676A-12743

Alignment Scores:
 Pred. No.: 3e-86 Length: 1839
 Score: 847.50 Matches: 162
 Percent Similarity: 97.63% Conservatives: 3
 Best Local Similarity: 95.86% Mismatches: 3
 Query Match: 81.26% Indels: 1
 DB: 5 Gaps: 1

US-09-782-953-7 (1-198) x US-09-724-676A-12743 (1-1839)
 Qy 30 AlaLysPheGluSerLeuPheArgThrTyraAspLysAspThrPheGlnTyPheLys 49
 Db 58 GCCAAATTTGAGTCCCTCTTTAGGACGTATGACAGGACATCACCTTTTCAGTATTTTAAAG 117
 Qy 50 SerPheLysArgValArgLleAsnPheserAsnProLysSerAlaAlaAspAlaArgLeu 69
 Db 118 AGCTTCAACAGGATCAGATAAACTTCAGCAACCCCTTCTCCGAGCAGATGCCAGGCTC 177

QY 70 ArgLeuHieLyThrGluPheLeuGlyLyGluMeuLyLeuTyPheAlaGlnThrLeu 89
 Db 178 CAGCTGCATACAGCTGAGTTCTGGGAAAGAAATGAATATATTTCTCAGACCTTA 237
 QY 90 HieLiegLySerSerHieLeuAlaProPheProAspPheGlnPheLeuLieserPro 109
 Db 238 CACATGAGAAAGCTCACACCTGGCTCCGCAAAATCCAGAACACAGTTCTGATCTCCCT 297
 QY 110 ProAlaSerProProValGlyTyrPheGlnValGluAspAlaThrProValLieserTy 129
 Db 298 CCGGCTCTCCGCGAGGTGGATGGAACAACTGGAAGATCCACCCAGTCAATTAACAT 357
 QY 130 AspLeuLeuTyAlaLieserLyLeuGlyProGlyGlyLeuTyGlyLeuHieAlaAla 149
 Db 358 GATCTCTTATATGCACTCTCCAAAGCTGGGCGCAGGGGAAAGATGAATTCACACGCG 417
 QY 150 ThrAspProThrProSerValValAlaHieValCysGluSerThrAspGlnLeuGluGlu 169
 Db 418 ACTGACACCACTCCAGCGTGGTGCATGATGAGAGATGATCMAAG--AAGGAG 474
 QY 170 GluGluGluGluMetGluArgMetLyAspGlyProLyPheLieserGlnThrArgArg 189
 Db 475 GAAGAAAGAGAAATGGAAGAAATGAGAGACCTTAAGCCAAAAATTCACAGACGAGAG 534
 QY 190 ProGluTyThrProLieserLeuSer 198
 Db 535 CCGAGTACACGCCGATCCACCTCAGC 561

RESULT 14

US-09-724-676-12749
 ; Sequence 12749, Application US/09724676
 ; GENERAL INFORMATION:
 ; APPLICANT: Compugen LTD
 ; TITLE OF INVENTION: Variants of alternative splicing
 ; FILE REFERENCE: 129181.4 Compugen
 ; CURRENT APPLICATION NUMBER: US/09/724,676
 ; CURRENT FILING DATE: 2000-11-28
 ; NUMBER OF SEQ ID NOS: 97222
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 12749
 ; LENGTH: 911
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (11)..(11)
 ; FEATURE:
 ; OTHER INFORMATION: n is a,c,g, or t
 ; NAME/KEY: misc_feature
 ; LOCATION: (18)..(18)
 ; OTHER INFORMATION: n is a,c,g, or t
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (23)..(23)
 ; OTHER INFORMATION: n is a,c,g, or t
 US-09-724-676-12749

Alignment Scores:

Pred. No.: 1,35e-70 Length: 911
 Score: 705.00 Matches: 133
 Percent Similarity: 97.20% Conservative: 6
 Best Local Similarity: 93.01% Mismatches: 4
 Query Match: 67.59% Indels: 0
 DB: 5 Gaps: 0

US-09-782-953-7 (1-198) x US-09-724-676-12749 (1-911)

QY 1 MetAspPheArgAspPheSerTyrrAsnPheserSerLeuLieserValAlaAsnAsp 20
 Db 63 ATGCAATTTTGAACCTTAACATGAGTTTACCTCCCTGATGCTGTGGCAACAGT 122
 QY 21 AspValPheSerGlySerGluThrArgAlaLysPheGluSerLeuPheArgThrTyrrAsp 40
 Db 123 GATATCTTCAGCGAAAGTGAACCAAGGCGCAAAATTTGAGTCCCTCTTTAGACGATATAC 182

Db 123 GATATCTTCAGCGAAAGTGAACCAAGGCGCAAAATTTGAGTCCCTCTTTAGACGATATAC 182
 QY 41 LysAspThrThrPheGlnTyrrPheLySerPheLyAspValArgLieserAsn 60
 Db 183 AAGACATACCTTCAGTATTTTAAAGAGCTTCAAGAGATCGAATTAATCTCAGAAC 242
 QY 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHieLysThrGluPheLeuGlyLyGlu 80
 Db 243 CCTCTTCGCGAGAGATGCGCAGGCTCAGTCAATGAGCTGTTCTGGGAAAGAA 302
 QY 81 MetLyLeuTyrrPheAlaGlnThrLeuHieLiegLySerSerHieLeuAlaProProAsn 100
 Db 303 ATGAGATTATATTTGCTCAGACCTTAACATGAGAGCTCACCTGGCTCCGCCAAAT 362
 QY 101 ProAspLyGlnPheLeuLieserProProAlaSerProProValGlyTyrPheGlnVal 120
 Db 363 CCAAGCAAGCAGTTTCTGATCTCCCTCCGCGCTCCGCGAGTGGAGAAACAAAGT 422
 QY 121 GluAspAlaThrProValLieserTyrrAspLeuTyrrAlaLieserLyLeuGlyPro 140
 Db 423 GAAGATGCGACCCAGATCATTAACATGATCTTATATGCAATCTCAAGCTGGGCGCA 482
 QY 141 GlyGlyLys 143
 Db 483 GGTGACGAG 491

RESULT 15

US-09-724-676A-12749
 ; Sequence 12749, Application US/09724676A
 ; GENERAL INFORMATION:
 ; APPLICANT: Compugen LTD
 ; TITLE OF INVENTION: Variants of alternative splicing
 ; FILE REFERENCE: 129181.4 Compugen
 ; CURRENT APPLICATION NUMBER: US/09/724,676A
 ; CURRENT FILING DATE: 2000-11-28
 ; NUMBER OF SEQ ID NOS: 97222
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 12749
 ; LENGTH: 911
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (11)..(11)
 ; OTHER INFORMATION: n is a,c,g, or t
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (18)..(18)
 ; OTHER INFORMATION: n is a,c,g, or t
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (23)..(23)
 ; OTHER INFORMATION: n is a,c,g, or t
 US-09-724-676A-12749

Alignment Scores:

Pred. No.: 1,35e-70 Length: 911
 Score: 705.00 Matches: 133
 Percent Similarity: 97.20% Conservative: 6
 Best Local Similarity: 93.01% Mismatches: 4
 Query Match: 67.59% Indels: 0
 DB: 5 Gaps: 0

US-09-782-953-7 (1-198) x US-09-724-676A-12749 (1-911)

QY 1 MetAspPheArgAspPheSerTyrrAsnPheserSerLeuLieserValAlaAsnAsp 20
 Db 63 ATGCAATTTTGAACCTTAACATGAGTTTACCTCCCTGATGCTGTGGCAACAGT 122
 QY 21 AspValPheSerGlySerGluThrArgAlaLysPheGluSerLeuPheArgThrTyrrAsp 40
 Db 123 GATATCTTCAGCGAAAGTGAACCAAGGCGCAAAATTTGAGTCCCTCTTTAGACGATATAC 182


```
Qy 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60
    |||||
Db 183 AAGGACATCACCTTTCAGTATTTTAAAGAGCTTCAAACGAGTTCAGATAAAACTTCAGCAAC 242
    |||||
Qy 61 ProLeuSerAlaAlaAspAlaAtgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
    |||||
Db 243 CCTTCTCCGAGCAGATGCCAGCTCCAGCTGCATTAAGACTGAGTTTCTGGGAAAGGAA 302
    |||||
Qy 81 MetLysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProProAsn 100
    |||||
Db 303 ATGAAGTTATATTTTGCTCAGACCTTACACATAGGAAGCTCACACCTGGCTCCGCCAAAT 362
    |||||
Qy 101 ProAspLysGlnPheLeuIleSerProProAlaSerProProValGlyTrpLysGlnVal 120
    |||||
Db 363 CCAGACAAGCAGTTTCTGATCTCCCTCCGCTCTCCGCGAGTGGGATGGAACAAGTG 422
    |||||
Qy 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro 140
    |||||
Db 423 GAAGATCGACCCCGAGTCATAACTATGATCTCTTATATGCCATCTCCAAAGCTGGGGCCA 482
    |||||
Qy 141 GlyGluLys 143
    |||||
Db 483 GGTGAGCAG 491
```

Search completed: December 15, 2002, 00:04:25
Job time : 57.7555 sec

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 11:39:49 ; Search time 6.34401 Seconds
(without alignments)
506.931 Million cell updates/sec

Title: US-09-782-953-7
Perfect score: 1043
Sequence: 1 MDRFDSYFNFSLLIACVAND.....RPKPKIIQTRRPEYTPHLS 198

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 103943 seqs, 16242309 residues

Total number of hits satisfying chosen parameters: 103943

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTU8_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1043	100.0	198	10	US-09-782-953-6
2	1043	100.0	198	10	US-09-782-953-7
3	970.5	93.0	197	10	US-09-782-953-12
4	970.5	93.0	197	10	US-09-782-953-13
5	920.5	88.3	198	10	US-09-782-953-3
6	920.5	88.3	198	10	US-09-782-953-4
7	875	83.9	197	10	US-09-782-953-15
8	875	83.9	197	10	US-09-782-953-16
9	632.5	60.6	197	10	US-09-782-953-9
10	632.5	60.6	197	10	US-09-782-953-10
11	615.5	59.0	192	10	US-09-782-953-18
12	615.5	59.0	192	10	US-09-782-953-19
13	592.5	56.8	241	10	US-09-782-953-21
14	592.5	56.8	241	10	US-09-782-953-22
15	591.5	56.7	212	10	US-09-782-953-24
16	591.5	56.7	212	10	US-09-782-953-25
17	369.5	35.4	142	10	US-09-925-302-790
18	284	27.2	58	10	US-09-864-761-43076
19	284	27.2	58	10	US-09-864-761-43076

20	271.5	26.0	56	10	US-09-864-761-34111	Sequence 34111, A
21	90.5	8.7	4019	10	US-09-738-973-425	Sequence 425, App
22	86.5	8.3	822	10	US-09-824-734-3	Sequence 3, Appli
23	81	7.8	2337	12	US-10-033-026-8	Sequence 8, Appli
24	81	7.8	2339	12	US-10-033-026-6	Sequence 6, Appli
25	81	7.8	2343	12	US-10-033-026-4	Sequence 4, Appli
26	79.5	7.6	559	10	US-09-854-549-7	Sequence 7, Appli
27	79.5	7.6	559	10	US-09-836-561-6	Sequence 6, Appli
28	77.5	7.4	864	10	US-09-883-096-2	Sequence 2, Appli
29	75.5	7.2	472	10	US-09-940-836A-2	Sequence 2, Appli
30	75	7.2	195	9	US-09-902-941-1888	Sequence 1888, Ap
31	75	7.2	362	9	US-09-349-755-32	Sequence 32, Appl
32	75	7.2	362	9	US-09-166-334-32	Sequence 32, Appl
33	75	7.2	362	10	US-09-350-206-32	Sequence 32, Appl
34	75	7.2	590	10	US-09-815-242-10812	Sequence 10812, A
35	75	7.2	1885	10	US-09-920-346-2	Sequence 2, Appli
36	74	7.1	311	10	US-09-819-252-2	Sequence 2, Appli
37	74	7.1	311	10	US-09-942-217-1061	Sequence 1061, Ap
38	74	7.1	311	10	US-09-833-263-1061	Sequence 1061, Ap
39	74	7.1	445	9	US-09-349-755-5	Sequence 5, Appli
40	74	7.1	445	9	US-09-166-334-5	Sequence 5, Appli
41	74	7.1	445	10	US-09-350-206-5	Sequence 5, Appli
42	74	7.1	916	10	US-09-745-763-174	Sequence 174, App
43	73	7.0	721	12	US-10-025-187-2	Sequence 2, Appli
44	72.5	7.0	1924	9	US-09-866-557A-2	Sequence 2, Appli
45	71.5	6.9	284	10	US-09-764-864-843	Sequence 843, App

ALIGNMENTS

RESULT 1
US-09-782-953-6
; Sequence 6, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674P21
; CURRENT APPLICATION NUMBER: US/09/782,953
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-782-953-6

Query Match	100.0%	Score	1043;	DB	10;	Length	198;
Best Local Similarity	100.0%	Pred. No.	4.5e-98;				
Matches	198;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	MDRFDPSYFNFSLLIACVANDDVSESETRAKESLFRTYDKDTTQYFKSKVRINFSN	60				
Db	1	MDRFDPSYFNFSLLIACVANDDVSESETRAKESLFRTYDKDTTQYFKSKVRINFSN	60				
Qy	61	PLSAADARLRLHKTFFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPAGPPVGVKQV	120				
Db	61	PLSAADARLRLHKTFFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPAGPPVGVKQV	120				
Qy	121	EDATPVINDLLYATSKLGPGEKYLHAATDTPSVVHVCSDDQENEEEMERMKRP	180				
Db	121	EDATPVINDLLYATSKLGPGEKYLHAATDTPSVVHVCSDDQENEEEMERMKRP	180				
Qy	181	KPKIIQTRRPEYTPHLS	198				
Db	181	KPKIIQTRRPEYTPHLS	198				

```
RESULT 2
US-09-782-953-7
; Sequence 7, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: US/09/782,953
; CURRENT APPLICATION NUMBER: US/09/782,953
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-782-953-7

Query Match          100.0%; Score 1043; DB 10; Length 198;
Best Local Similarity 100.0%; Pred. No. 4,5e-98;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 MDRFDSYNSFSLIACVANDVFSSESETRAKFESLFRITDKDTTFQYFKSPKVRINFSN 60
Db 1 MDRFDSYNSFSLIACVANDVFSSESETRAKFESLFRITDKDTTFQYFKSPKVRINFSN 60
Oy 61 PLSAADARLRHLKTEFLGKEMKLYFAQTLHIGSSHLAPNPKQFLISPPASPVGMKQV 120
Db 61 PLSAADARLRHLKTEFLGKEMKLYFAQTLHIGSSHLAPNPKQFLISPPASPVGMKQV 120
Oy 121 EDATPVINVDLLVAISKLGPEKYEHLAATDPTPSVVVHVCSQDENEEEMERMRKP 180
Db 121 EDATPVINVDLLVAISKLGPEKYEHLAATDPTPSVVVHVCSQDENEEEMERMRKP 180
Oy 181 KPRTIOTRRPEYTPILHS 198
Db 181 KPRTIOTRRPEYTPILHS 198

RESULT 3
US-09-782-953-12
; Sequence 12, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: US/09/782,953
; CURRENT APPLICATION NUMBER: US/09/782,953
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-12

Query Match          93.0%; Score 970.5; DB 10; Length 197;
Best Local Similarity 93.4%; Pred. No. 9,2e-91;
Matches 185; Conservative 7; Mismatches 5; Indels 1; Gaps 1;

Oy 1 MDRFDSYNSFSLIACVANDVFSSESETRAKFESLFRITDKDTTFQYFKSPKVRINFSN 60
Db 1 MDRFDSYNSFSLIACVANDVFSSESETRAKFESLFRITDKDTTFQYFKSPKVRINFSN 60
```

```
Oy 61 PLSAADARLRHLKTEFLGKEMKLYFAQTLHIGSSHLAPNPKQFLISPPASPVGMKQV 120
Db 61 PLSAADARLRHLKTEFLGKEMKLYFAQTLHIGSSHLAPNPKQFLISPPASPVGMKQV 120
Oy 121 EDATPVINVDLLVAISKLGPEKYEHLAATDPTPSVVVHVCSQDENEEEMERMRKP 180
Db 121 EDATPVINVDLLVAISKLGPEKYEHLAATDPTPSVVVHVCSQDENEEEMERMRKP 180
Oy 181 KPRTIOTRRPEYTPILHS 198
Db 181 KPRTIOTRRPEYTPILHS 197

RESULT 4
US-09-782-953-13
; Sequence 13, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: US/09/782,953
; CURRENT APPLICATION NUMBER: US/09/782,953
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-13

Query Match          93.0%; Score 970.5; DB 10; Length 197;
Best Local Similarity 93.4%; Pred. No. 9,2e-91;
Matches 185; Conservative 7; Mismatches 5; Indels 1; Gaps 1;

Oy 1 MDRFDSYNSFSLIACVANDVFSSESETRAKFESLFRITDKDTTFQYFKSPKVRINFSN 60
Db 1 MDRFDSYNSFSLIACVANDVFSSESETRAKFESLFRITDKDTTFQYFKSPKVRINFSN 60
Oy 61 PLSAADARLRHLKTEFLGKEMKLYFAQTLHIGSSHLAPNPKQFLISPPASPVGMKQV 120
Db 61 PLSAADARLRHLKTEFLGKEMKLYFAQTLHIGSSHLAPNPKQFLISPPASPVGMKQV 120
Oy 121 EDATPVINVDLLVAISKLGPEKYEHLAATDPTPSVVVHVCSQDENEEEMERMRKP 180
Db 121 EDATPVINVDLLVAISKLGPEKYEHLAATDPTPSVVVHVCSQDENEEEMERMRKP 179
Oy 181 KPRTIOTRRPEYTPILHS 198
Db 181 KPRTIOTRRPEYTPILHS 197

RESULT 5
US-09-782-953-3
; Sequence 3, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: US/09/782,953
; CURRENT APPLICATION NUMBER: US/09/782,953
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
```


QY 121 EDATVINVLDLYAISKLGPGKXELHAATDPTPSVVVHVCSDDQNEEBEEMERMRKP 180
 |||
 Db 121 EDATVINVLDLYAISKLGPGKXELHAATDPTPSVVVHVCSDDQ-KEEBEEMERMRP 179
 |||
 QY 181 KPKIOTRRPEYTPIHLS 198
 |||
 Db 180 KPKIOTRRPEYTPIHLS 197
 |||

RESULT 9

US-09-782-953-9
 ; Sequence 9, Application US/09782953
 ; Patent No. US20020150953A1
 ; GENERAL INFORMATION:
 ; APPLICANT: WILLIAMS, R. SANDERS
 ; APPLICANT: ROTHERMEL, BEVERLY
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
 ; TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)
 ; FILE REFERENCE: US/09/782,953
 ; CURRENT FILING DATE: 2001-02-13
 ; PRIOR APPLICATION NUMBER: 60/216,601
 ; PRIOR FILING DATE: 2000-07-07
 ; NUMBER OF SEQ ID NOS: 27
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 9
 ; LENGTH: 197
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 US-09-782-953-9

Query Match 60.6%; Score 632.5; DB 10; Length 197;
 Best Local Similarity 64.4%; Pred. No. 1.2e-56;
 Matches 121; Conservative 24; Mismatches 34; Indels 9; Gaps 2;

QY 11 SGLIACVANDVFSSESETRAKFESLFRYYDKDTTFQYFKSFKRVIRINFSNPLSADARLR 70
 |||
 Db 11 STLVACVVDVEFTNOEVEKEFGGLFRYYDCCVTFQLFKSPFRVRINFSHPSAARARIE 70
 |||
 QY 71 LKTEFLGKEMKLYFAQTLHIGSS----HLAPNPDKQFLISPPASPPVGMKQVEDATPV 126
 |||
 Db 71 LHEQFRGKKLKYFAQVQTFETDGDKHLAPQPAKQFLISPPSSPPVGMKPISDATPV 130
 |||
 QY 127 INVLDLYAISKLGPGKXELHAATDPTPSVVVHVCSDDQNEEBEEMERMRKP 186
 |||
 Db 131 INVLDLYAISKLGPGKXELHAATDPTPSVVVHVCSDDQNEEBEEMERMRKP 185
 |||
 QY 187 TRRPEYTP 194
 |||
 Db 186 TRRPEYTP 193
 |||

RESULT 10

US-09-782-953-10
 ; Sequence 10, Application US/09782953
 ; Patent No. US20020150953A1
 ; GENERAL INFORMATION:
 ; APPLICANT: WILLIAMS, R. SANDERS
 ; APPLICANT: ROTHERMEL, BEVERLY
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
 ; TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)
 ; FILE REFERENCE: US/09/782,953
 ; CURRENT FILING DATE: 2001-02-13
 ; PRIOR APPLICATION NUMBER: 60/216,601
 ; PRIOR FILING DATE: 2000-07-07
 ; NUMBER OF SEQ ID NOS: 27
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 10
 ; LENGTH: 197
 ; TYPE: PRT
 ; ORGANISM: Mus musculus

US-09-782-953-10

Query Match 60.6%; Score 632.5; DB 10; Length 197;
 Best Local Similarity 64.4%; Pred. No. 1.2e-56;
 Matches 121; Conservative 24; Mismatches 34; Indels 9; Gaps 2;

QY 11 SGLIACVANDVFSSESETRAKFESLFRYYDKDTTFQYFKSFKRVIRINFSNPLSADARLR 70
 |||
 Db 11 STLVACVVDVEFTNOEVEKEFGGLFRYYDCCVTFQLFKSPFRVRINFSHPSAARARIE 70
 |||
 QY 71 LKTEFLGKEMKLYFAQTLHIGSS----HLAPNPDKQFLISPPASPPVGMKQVEDATPV 126
 |||
 Db 71 LHEQFRGKKLKYFAQVQTFETDGDKHLAPQPAKQFLISPPSSPPVGMKPISDATPV 130
 |||
 QY 127 INVLDLYAISKLGPGKXELHAATDPTPSVVVHVCSDDQNEEBEEMERMRKP 186
 |||
 Db 131 INVLDLYAISKLGPGKXELHAATDPTPSVVVHVCSDDQNEEBEEMERMRKP 185
 |||
 QY 187 TRRPEYTP 194
 |||
 Db 186 TRRPEYTP 193
 |||

RESULT 11

US-09-782-953-18
 ; Sequence 18, Application US/09782953
 ; Patent No. US20020150953A1
 ; GENERAL INFORMATION:
 ; APPLICANT: WILLIAMS, R. SANDERS
 ; APPLICANT: ROTHERMEL, BEVERLY
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
 ; TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)
 ; FILE REFERENCE: US/09/782,953
 ; CURRENT FILING DATE: 2001-02-13
 ; PRIOR APPLICATION NUMBER: 60/216,601
 ; PRIOR FILING DATE: 2000-07-07
 ; NUMBER OF SEQ ID NOS: 27
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 18
 ; LENGTH: 192
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-782-953-18

Query Match 59.0%; Score 615.5; DB 10; Length 192;
 Best Local Similarity 63.3%; Pred. No. 6e-55;
 Matches 119; Conservative 23; Mismatches 37; Indels 9; Gaps 2;

QY 11 SGLIACVANDVFSSESETRAKFESLFRYYDKDTTFQYFKSFKRVIRINFSNPLSADARLR 70
 |||
 Db 6 STLVACVVDVEFTNOEVEKEFGGLFRYYDCCVTFQLFKSPFRVRINFSHPSAARARIE 65
 |||
 QY 71 LKTEFLGKEMKLYFAQTLHIGSS----HLAPNPDKQFLISPPASPPVGMKQVEDATPV 126
 |||
 Db 66 LHEQFRGKKLKYFAQVQTFETDGDKHLAPQPAKQFLISPPSSPPVGMKPISDATPV 125
 |||
 QY 127 INVLDLYAISKLGPGKXELHAATDPTPSVVVHVCSDDQNEEBEEMERMRKP 186
 |||
 Db 126 INVLDLYAISKLGPGKXELHAATDPTPSVVVHVCSDDQNEEBEEMERMRKP 180
 |||
 QY 187 TRRPEYTP 194
 |||
 Db 181 TRRPEYTP 188
 |||

RESULT 12

US-09-782-953-19
 ; Sequence 19, Application US/09782953
 ; Patent No. US20020150953A1
 ; GENERAL INFORMATION:
 ; APPLICANT: WILLIAMS, R. SANDERS
 ; APPLICANT: ROTHERMEL, BEVERLY

```
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674P21
; CURRENT APPLICATION NUMBER: US/09/782,953
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-19

Query Match          59.0%; Score 615.5; DB 10; Length 192;
Best Local Similarity 63.3%; Pred. No. 6e-55;
Matches 119; Conservative 23; Mismatches 37; Indels 9; Gaps 2;

QY 11 SLLIACVANDVFSESETRAKFESLRTYDKDTTFQYFKSFKRVRINFNSPLSADARLR 70
DB 6 STLVACVQDVVEFTNOEVEKFGGLRTYDDCVTFQYFKSFRRVRINFNSPKSARARIE 65
QY 71 LKHTFLGKEMKLYFAQTLHIGSS-----HLAPNPDKQLISPPASPPVGMKQVEDATPV 126
DB 66 LHETQGRGKLLKLYFAQVQTPETDQGLHLAPPQPAKQQLISPPSPVPVSMQINDATPV 125
QY 127 INYDLLYALSKLPGKGYELHAATDTPSVVHVHVCSDQENEEEMERMKRPKPIIQ 186
DB 126 LNYDLLYAVAKLPGKGYELHAGTSTPSVVHVHVCSDIEEDDPK-----TSPKPKIIQ 180
QY 187 TRPEYTP 194
DB 181 TRPGLPP 188

RESULT 13
US-09-782-953-21
; Sequence 21, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674P21
; CURRENT APPLICATION NUMBER: US/09/782,953
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-21

Query Match          56.8%; Score 592.5; DB 10; Length 241;
Best Local Similarity 61.5%; Pred. No. 1.7e-52;
Matches 120; Conservative 19; Mismatches 43; Indels 13; Gaps 3;

QY 1 MDRDFSYNFFSLIACVANDVFSESETRAKFESLRTYDKDTTFQYFKSFKRVRINFNS 60
DB 40 MDLSLP---TSLFACSVHEAVFEAREQERFEALTIYDDQVTFQYFKSFRRVRINF 96
QY 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHIG-----SSHAPPNPDKQLISPPASPPVG 116
DB 97 PEAARARIELHETDFNGOKLKYFAQVQMSGEVRDKSYLLPPQPVKQFLISPPASPPVG 156
QY 117 WKQVEDATPVINYDLLYALSKLPGKGYELHAATDTPSVVHVHVCSDQENEEEMER 176
DB 157 WKQSDAMPVINYDLLCAVSKLPGKGYELHAGTSTPSVVHVHVCSETEEE----- 210

RESULT 15
US-09-782-953-24
; Sequence 24, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674P21
; CURRENT APPLICATION NUMBER: US/09/782,953
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 24
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-24

Query Match          56.7%; Score 591.5; DB 10; Length 212;
Best Local Similarity 62.3%; Pred. No. 1.8e-52;
```

```
QY 177 MKRPKPKIIQTRRPE 191
DB 211 TKNPKQKIAQTRRPD 225

RESULT 14
US-09-782-953-22
; Sequence 22, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674P21
; CURRENT APPLICATION NUMBER: US/09/782,953
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-22

Query Match          56.8%; Score 592.5; DB 10; Length 241;
Best Local Similarity 61.5%; Pred. No. 1.7e-52;
Matches 120; Conservative 19; Mismatches 43; Indels 13; Gaps 3;

QY 1 MDRDFSYNFFSLIACVANDVFSESETRAKFESLRTYDKDTTFQYFKSFKRVRINFNS 60
DB 40 MDLSLP---TSLFACSVHEAVFEAREQERFEALTIYDDQVTFQYFKSFRRVRINF 96
QY 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHIG-----SSHAPPNPDKQLISPPASPPVG 116
DB 97 PEAARARIELHETDFNGOKLKYFAQVQMSGEVRDKSYLLPPQPVKQFLISPPASPPVG 156
QY 117 WKQVEDATPVINYDLLYALSKLPGKGYELHAATDTPSVVHVHVCSDQENEEEMER 176
DB 157 WKQSDAMPVINYDLLCAVSKLPGKGYELHAGTSTPSVVHVHVCSETEEE----- 210

RESULT 15
US-09-782-953-24
; Sequence 24, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674P21
; CURRENT APPLICATION NUMBER: US/09/782,953
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 24
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-24

Query Match          56.7%; Score 591.5; DB 10; Length 212;
Best Local Similarity 62.3%; Pred. No. 1.8e-52;
```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 11:38:44 ; Search time 136.563 Seconds
(without alignments)
934.784 Million cell updates/sec

Title: US-09-782-953-7

Perfect score: 1043

Sequence: 1 MDRDFSYNFSLLIACVAND.....RKPKEIQTRRPEYTHLS 198

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents AA Main.*
1: /cgn2_6/ptodata/1/paa/PCTUS COMB.pcp.*
2: /cgn2_6/ptodata/1/paa/US06 COMB.pcp.*
3: /cgn2_6/ptodata/1/paa/US07 COMB.pcp.*
4: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
5: /cgn2_6/ptodata/1/paa/US081 COMB.pcp.*
6: /cgn2_6/ptodata/1/paa/US082 COMB.pcp.*
7: /cgn2_6/ptodata/1/paa/US083 COMB.pcp.*
8: /cgn2_6/ptodata/1/paa/US084 COMB.pcp.*
9: /cgn2_6/ptodata/1/paa/US085 COMB.pcp.*
10: /cgn2_6/ptodata/1/paa/US086 COMB.pcp.*
11: /cgn2_6/ptodata/1/paa/US087 COMB.pcp.*
12: /cgn2_6/ptodata/1/paa/US088 COMB.pcp.*
13: /cgn2_6/ptodata/1/paa/US089 COMB.pcp.*
14: /cgn2_6/ptodata/1/paa/US090 COMB.pcp.*
15: /cgn2_6/ptodata/1/paa/US091 COMB.pcp.*
16: /cgn2_6/ptodata/1/paa/US092 COMB.pcp.*
17: /cgn2_6/ptodata/1/paa/US093 COMB.pcp.*
18: /cgn2_6/ptodata/1/paa/US094 COMB.pcp.*
19: /cgn2_6/ptodata/1/paa/US095 COMB.pcp.*
20: /cgn2_6/ptodata/1/paa/US096 COMB.pcp.*
21: /cgn2_6/ptodata/1/paa/US097 COMB.pcp.*
22: /cgn2_6/ptodata/1/paa/US098 COMB.pcp.*
23: /cgn2_6/ptodata/1/paa/US099 COMB.pcp.*
24: /cgn2_6/ptodata/1/paa/US100 COMB.pcp.*
25: /cgn2_6/ptodata/1/paa/US101 COMB.pcp.*
26: /cgn2_6/ptodata/1/paa/US102 COMB.pcp.*
27: /cgn2_6/ptodata/1/paa/US60 COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1043	100.0	198	21	US-09-782-953-6
2	1043	100.0	198	21	US-09-782-953-7
3	992.5	95.2	197	19	US-09-575-580B-8
4	970.5	93.0	197	21	US-09-782-953-12
5	970.5	93.0	197	21	US-09-782-953-13
6	970.5	93.0	197	26	US-10-247-671-174

7	970.5	93.0	197	27	US-60-323-784-174	Sequence 174, Appl
8	920.5	88.3	198	19	US-09-575-580B-4	Sequence 4, Appl
9	920.5	88.3	198	21	US-09-782-953-3	Sequence 3, Appl
10	920.5	88.3	198	21	US-09-782-953-4	Sequence 4, Appl
11	875	83.9	197	21	US-09-782-953-15	Sequence 15, Appl
12	875	83.9	197	21	US-09-782-953-16	Sequence 16, Appl
13	875	83.9	197	21	US-10-104-047-2892	Sequence 16, Appl
14	843.5	80.9	197	21	US-09-575-580B-7	Sequence 2892, Ap
15	632.5	60.6	170	19	US-09-782-953-9	Sequence 9, Appl
16	632.5	60.6	197	21	US-09-782-953-10	Sequence 10, Appl
17	628.5	60.3	192	19	US-09-575-580B-6	Sequence 6, Appl
18	628.5	60.3	197	1	PCT-US02-17382-132	Sequence 132, App
19	623	59.7	234	24	US-10-030-613-1	Sequence 1, Appl
20	623	59.7	234	27	US-60-142-678-1	Sequence 1, Appl
21	623	59.7	243	25	US-10-104-047-2216	Sequence 2216, Ap
22	623	59.7	255	20	US-09-614-474-2	Sequence 2, Appl
23	615.5	59.0	192	20	US-09-614-474-10	Sequence 10, Appl
24	615.5	59.0	192	21	US-09-782-953-18	Sequence 18, Appl
25	615.5	59.0	192	21	US-09-782-953-19	Sequence 19, Appl
26	613.5	58.8	242	19	US-09-575-580B-5	Sequence 5, Appl
27	592.5	56.8	241	1	PCT-US02-17382-133	Sequence 133, App
28	592.5	56.8	241	20	US-09-614-474-11	Sequence 11, Appl
29	592.5	56.8	241	21	US-09-782-953-21	Sequence 21, Appl
30	592.5	56.8	241	21	US-09-782-953-22	Sequence 22, Appl
31	591.5	56.7	212	21	US-09-782-953-24	Sequence 24, Appl
32	591.5	56.7	212	21	US-09-782-953-25	Sequence 25, Appl
33	581	55.7	236	19	US-09-575-580B-24	Sequence 24, Appl
34	432	41.4	142	21	US-09-758-472-6466	Sequence 6466, Ap
35	432	41.4	142	26	US-10-235-926-6466	Sequence 6466, Ap
36	372	35.7	292	20	US-09-614-150-41193	Sequence 41193, A
37	372	35.7	292	20	US-09-619-049-1440	Sequence 1440, Ap
38	372	35.7	292	27	US-60-167-324-718	Sequence 718, App
39	372	35.7	292	27	US-60-171-627-2109	Sequence 2109, Ap
40	372	35.7	292	27	US-60-173-386-680	Sequence 680, App
41	372	35.7	292	27	US-60-175-871-764	Sequence 764, App
42	372	35.7	292	27	US-60-184-775-692	Sequence 692, App
43	372	35.7	292	27	US-60-191-637-40818	Sequence 40818, A
44	372	35.7	292	27	US-60-191-700-744	Sequence 744, App
45	369.5	35.4	142	1	PCT-US00-05918-790	Sequence 790, App

ALIGNMENTS

RESULT 1
US-09-782-953-6
; Sequence 6, Application US/09782953
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTS02:674E21
; CURRENT APPLICATION NUMBER: US/09/782,953
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-782-953-6

Query Match 100.0%; Score 1043; DB 21; Length 198;
Best Local Similarity 100.0%; Pred. No. 4.1e-95;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MDRDFSYNFSLLIACVANDVFSESETRAKFESLFRTYDKDTTFQYKFKRVRINFSN 60
Db 1 MDRDFSYNFSLLIACVANDVFSESETRAKFESLFRTYDKDTTFQYKFKRVRINFSN 60

QY 61 PLSAADARLRLHKTFFLGKEMKLYPAQTLLHIGSSHLAPNPDKQFLISPPASPPVGMKQV 120
 DB 61 PLSAADARLRLHKTFFLGKEMKLYPAQTLLHIGSSHLAPNPDKQFLISPPASPPVGMKQV 120
 QY 121 EDATPVINVDLLVAISKLGEGEYELHAATDPTPSVVVHVCSDOENEBEEMERMKRP 180
 DB 121 EDATPVINVDLLVAISKLGEGEYELHAATDPTPSVVVHVCSDOENEBEEMERMKRP 180
 QY 181 KPRTIOTRRPEYTPHLS 198
 DB 181 KPRTIOTRRPEYTPHLS 198

RESULT 2
 US-09-782-953-7
 ; Sequence 7, Application US/09782953
 ; GENERAL INFORMATION:
 ; APPLICANT: WILLIAMS, R. SANDERS
 ; APPLICANT: ROTHERMEL, BEVERLY
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
 ; FILE REFERENCE: UTSD:674P21
 ; CURRENT APPLICATION NUMBER: US/09/782,953
 ; PRIOR FILING DATE: 2001-02-13
 ; PRIOR APPLICATION NUMBER: 60/216,601
 ; NUMBER OF SEQ ID NOS: 27
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 7
 ; LENGTH: 198
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 US-09-782-953-7

Query Match 100.0%; Score 1043; DB 21; Length 198;
 Best Local Similarity 100.0%; Pred. No. 4.1e-95;
 Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDRDPSYNFSSLIACVANDVSESETRAKFESLFRYYDKOTTFQYFKSPKVRINFSN 60
 DB 1 MDRDPSYNFSSLIACVANDVSESETRAKFESLFRYYDKOTTFQYFKSPKVRINFSN 60
 QY 61 PLSAADARLRLHKTFFLGKEMKLYPAQTLLHIGSSHLAPNPDKQFLISPPASPPVGMKQV 120
 DB 61 PLSAADARLRLHKTFFLGKEMKLYPAQTLLHIGSSHLAPNPDKQFLISPPASPPVGMKQV 120
 QY 121 EDATPVINVDLLVAISKLGEGEYELHAATDPTPSVVVHVCSDOENEBEEMERMKRP 180
 DB 121 EDATPVINVDLLVAISKLGEGEYELHAATDPTPSVVVHVCSDOENEBEEMERMKRP 180
 QY 181 KPRTIOTRRPEYTPHLS 198
 DB 181 KPRTIOTRRPEYTPHLS 198

RESULT 3
 US-09-575-5808-8
 ; Sequence 8, Application US/095755808
 ; GENERAL INFORMATION:
 ; APPLICANT: MCKEON, P.
 ; APPLICANT: KAYAKO, K.
 ; APPLICANT: RYGEOM, S.
 ; TITLE OF INVENTION: CALCIPRESSINS: ENDOGENOUS INHIBITORS OF CALCINEURIN
 ; TITLE OF INVENTION: USES AND REAGENTS RELATED THERETO
 ; FILE REFERENCE: HMV-048.01
 ; CURRENT APPLICATION NUMBER: US/09/575,580B
 ; NUMBER OF SEQ ID NOS: 45
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 8
 ; LENGTH: 197
 ; TYPE: PRT
 ; ORGANISM: Citicellus griseus

US-09-575-5808-8

Query Match 95.2%; Score 992.5; DB 19; Length 197;
 Best Local Similarity 96.5%; Pred. No. 4.4e-90;
 Matches 191; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 1 MDRDPSYNFSSLIACVANDVSESETRAKFESLFRYYDKOTTFQYFKSPKVRINFSN 60
 DB 1 MDRDPSYNFSSLIACVANDVSESETRAKFESLFRYYDKOTTFQYFKSPKVRINFSN 60
 QY 61 PLSAADARLRLHKTFFLGKEMKLYPAQTLLHIGSSHLAPNPDKQFLISPPASPPVGMKQV 120
 DB 61 PLSAADARLRLHKTFFLGKEMKLYPAQTLLHIGSSHLAPNPDKQFLISPPASPPVGMKQV 120
 QY 121 EDATPVINVDLLVAISKLGEGEYELHAATDPTPSVVVHVCSDOENEBEEMERMKRP 180
 DB 121 EDATPVINVDLLVAISKLGEGEYELHAATDPTPSVVVHVCSDOENEBEEMERMKRP 179
 QY 181 KPRTIOTRRPEYTPHLS 198
 DB 180 KPRTIOTRRPEYTPHLS 197

RESULT 4
 US-09-782-953-12
 ; Sequence 12, Application US/09782953
 ; GENERAL INFORMATION:
 ; APPLICANT: WILLIAMS, R. SANDERS
 ; APPLICANT: ROTHERMEL, BEVERLY
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
 ; FILE REFERENCE: UTSD:674P21
 ; CURRENT APPLICATION NUMBER: US/09/782,953
 ; PRIOR FILING DATE: 2001-02-13
 ; PRIOR APPLICATION NUMBER: 60/216,601
 ; NUMBER OF SEQ ID NOS: 27
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 12
 ; LENGTH: 197
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-782-953-12

Query Match 93.0%; Score 970.5; DB 21; Length 197;
 Best Local Similarity 93.4%; Pred. No. 6.9e-86;
 Matches 185; Conservative 7; Mismatches 5; Indels 1; Gaps 1;

QY 1 MDRDPSYNFSSLIACVANDVSESETRAKFESLFRYYDKOTTFQYFKSPKVRINFSN 60
 DB 1 MDRDPSYNFSSLIACVANDVSESETRAKFESLFRYYDKOTTFQYFKSPKVRINFSN 60
 QY 61 PLSAADARLRLHKTFFLGKEMKLYPAQTLLHIGSSHLAPNPDKQFLISPPASPPVGMKQV 120
 DB 61 PLSAADARLRLHKTFFLGKEMKLYPAQTLLHIGSSHLAPNPDKQFLISPPASPPVGMKQV 120
 QY 121 EDATPVINVDLLVAISKLGEGEYELHAATDPTPSVVVHVCSDOENEBEEMERMKRP 180
 DB 121 EDATPVINVDLLVAISKLGEGEYELHAATDPTPSVVVHVCSDOENEBEEMERMKRP 179
 QY 181 KPRTIOTRRPEYTPHLS 198
 DB 180 KPRTIOTRRPEYTPHLS 197

RESULT 5
 US-09-782-953-13
 ; Sequence 13, Application US/09782953
 ; GENERAL INFORMATION:
 ; APPLICANT: WILLIAMS, R. SANDERS
 ; APPLICANT: ROTHERMEL, BEVERLY
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
 ; TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)

```
; FILE REFERENCE: UTSD.674P21
; CURRENT APPLICATION NUMBER: US/09/782,953
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-13

Query Match      93.0%; Score 970.5; DB 21; Length 197;
Best Local Similarity 93.4%; Pred. No. 6.9e-88;
Matches 185; Conservative 7; Mismatches 5; Indels 1; Gaps 1;

QY 1 MDRFDSYFSSLIACVANDVSESETRAKFESLFRDYDKDTTFQYFKSKRVINFSN 60
Db 1 MHFRNFYFSSLIACVANDVSESETRAKFESLFRDYDKDTTFQYFKSKRVINFSN 60

QY 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPPVGWQV 120
Db 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPPVGWQV 120

QY 121 EDATPVINYDLLYAIKLGPGKVELHAATDPTPSVVHVHVCESDQENEEBEMERMRP 180
Db 121 EDATPVINYDLLYAIKLGPGKVELHAATDPTPSVVHVHVCESDQENEEBEMERMRP 180

QY 181 KPKIOTRRPEYTPHLS 198
Db 180 KPKIOTRRPEYTPHLS 197

RESULT 6
US-10-247-671-174
; Sequence 174, Application US/10247671
; GENERAL INFORMATION:
; APPLICANT: Mikita, Thomas
; APPLICANT: Shiffman, Dov
; APPLICANT: Porter, Gordon, J.
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS
; FILE REFERENCE: PA-0050 US
; CURRENT APPLICATION NUMBER: US/10/247,671
; CURRENT FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/323,784
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: PERL Program
; SEQ ID NO 174
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 5511889CD1
US-10-247-671-174

Query Match      93.0%; Score 970.5; DB 26; Length 197;
Best Local Similarity 93.4%; Pred. No. 6.9e-88;
Matches 185; Conservative 7; Mismatches 5; Indels 1; Gaps 1;

QY 1 MDRFDSYFSSLIACVANDVSESETRAKFESLFRDYDKDTTFQYFKSKRVINFSN 60
Db 1 MHFRNFYFSSLIACVANDVSESETRAKFESLFRDYDKDTTFQYFKSKRVINFSN 60

QY 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPPVGWQV 120
Db 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPPVGWQV 120

QY 121 EDATPVINYDLLYAIKLGPGKVELHAATDPTPSVVHVHVCESDQENEEBEMERMRP 180
Db 121 EDATPVINYDLLYAIKLGPGKVELHAATDPTPSVVHVHVCESDQENEEBEMERMRP 180

QY 181 KPKIOTRRPEYTPHLS 198
Db 180 KPKIOTRRPEYTPHLS 197

RESULT 7
US-60-323-784-174
; Sequence 174, Application US/60323784
; GENERAL INFORMATION:
; APPLICANT: Mikita, Thomas
; APPLICANT: Shiffman, Dov
; APPLICANT: Porter, Gordon, J.
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS
; FILE REFERENCE: PA-0050 P
; CURRENT APPLICATION NUMBER: US/60/323,784
; CURRENT FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: PERL Program
; SEQ ID NO 174
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 5511889CD1
US-60-323-784-174

Query Match      93.0%; Score 970.5; DB 27; Length 197;
Best Local Similarity 93.4%; Pred. No. 6.9e-88;
Matches 185; Conservative 7; Mismatches 5; Indels 1; Gaps 1;

QY 1 MDRFDSYFSSLIACVANDVSESETRAKFESLFRDYDKDTTFQYFKSKRVINFSN 60
Db 1 MHFRNFYFSSLIACVANDVSESETRAKFESLFRDYDKDTTFQYFKSKRVINFSN 60

QY 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPPVGWQV 120
Db 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPPVGWQV 120

QY 121 EDATPVINYDLLYAIKLGPGKVELHAATDPTPSVVHVHVCESDQENEEBEMERMRP 180
Db 121 EDATPVINYDLLYAIKLGPGKVELHAATDPTPSVVHVHVCESDQENEEBEMERMRP 180

QY 181 KPKIOTRRPEYTPHLS 198
Db 180 KPKIOTRRPEYTPHLS 197

RESULT 8
US-09-575-580B-4
; Sequence 4, Application US/09575580B
; GENERAL INFORMATION:
; APPLICANT: McKeon, F.
; APPLICANT: Kayako, K.
; APPLICANT: Ryeom, S.
; TITLE OF INVENTION: CALCIPRESSINS: ENDOGENOUS INHIBITORS OF CALCINEURIN
; FILE REFERENCE: HMV-048.01
; CURRENT APPLICATION NUMBER: US/09/575,580B
; CURRENT FILING DATE: 2000-05-22
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-575-580B-4

Query Match      88.3%; Score 920.5; DB 19; Length 198;
Best Local Similarity 89.9%; Pred. No. 6.7e-83;
```

Matches 178; Conservative 5; Mismatches 12; Indels 3; Gaps 1;

```

QY 1 MDRPDSYNSFLICVANDVSESETRAKFESLFRITYDKOTTFQYKSFKRVINFSN 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 4 VDLQDLP--SATIACHLDPRVFDGLCRAKFESLFRITYDKOTTFQYKSFKRVINFSN 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 61 PLSAADARLRLHKTETFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPPVGMKQV 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 PLSAADARLRLHKTETFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPPVGMKQV 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 121 EDATPVINVDLLVAISKLGPEKYTELHAATDTPPSVVHVHCSDQENEEBEMERMRKP 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 121 EDATPVINVDLLVAISKLGPEKYTELHAATDTPPSVVHVHCSDQENEEBEMERMRKP 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 181 KPKIOTRRPEYTPHLS 198
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 181 KPKIOTRRPEYTPHLS 198
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 9

```

US-09-782-953-3
; Sequence 3, Application US/09782953
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674P21
; CURRENT APPLICATION NUMBER: US/09/782,953
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-782-953-3

```

Query Match 88.3%; Score 920.5; DB 21; Length 198;

Best Local Similarity 89.9%; Pred. No. 6.7e-83; Indels 3; Gaps 1;

Matches 178; Conservative 5; Mismatches 12; Indels 3; Gaps 1;

```

QY 1 MDRPDSYNSFLICVANDVSESETRAKFESLFRITYDKOTTFQYKSFKRVINFSN 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 4 VDLQDLP--SATIACHLDPRVFDGLCRAKFESLFRITYDKOTTFQYKSFKRVINFSN 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 61 PLSAADARLRLHKTETFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPPVGMKQV 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 PLSAADARLRLHKTETFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPPVGMKQV 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 121 EDATPVINVDLLVAISKLGPEKYTELHAATDTPPSVVHVHCSDQENEEBEMERMRKP 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 121 EDATPVINVDLLVAISKLGPEKYTELHAATDTPPSVVHVHCSDQENEEBEMERMRKP 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 181 KPKIOTRRPEYTPHLS 198
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 181 KPKIOTRRPEYTPHLS 198
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 10

```

US-09-782-953-4
; Sequence 4, Application US/09782953
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674P21
; CURRENT APPLICATION NUMBER: US/09/782,953
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601

```

```

; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-782-953-4

```

Query Match 88.3%; Score 920.5; DB 21; Length 198;

Best Local Similarity 89.9%; Pred. No. 6.7e-83; Indels 3; Gaps 1;

Matches 178; Conservative 5; Mismatches 12; Indels 3; Gaps 1;

```

QY 1 MDRPDSYNSFLICVANDVSESETRAKFESLFRITYDKOTTFQYKSFKRVINFSN 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 4 VDLQDLP--SATIACHLDPRVFDGLCRAKFESLFRITYDKOTTFQYKSFKRVINFSN 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 61 PLSAADARLRLHKTETFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPPVGMKQV 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 PLSAADARLRLHKTETFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPPVGMKQV 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 121 EDATPVINVDLLVAISKLGPEKYTELHAATDTPPSVVHVHCSDQENEEBEMERMRKP 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 121 EDATPVINVDLLVAISKLGPEKYTELHAATDTPPSVVHVHCSDQENEEBEMERMRKP 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 181 KPKIOTRRPEYTPHLS 198
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 181 KPKIOTRRPEYTPHLS 198
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 11

```

US-09-782-953-15
; Sequence 15, Application US/09782953
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674P21
; CURRENT APPLICATION NUMBER: US/09/782,953
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-15

```

Query Match 83.9%; Score 875; DB 21; Length 197;

Best Local Similarity 86.4%; Pred. No. 2.3e-78; Indels 4; Gaps 2;

Matches 171; Conservative 8; Mismatches 15; Indels 4; Gaps 2;

```

QY 1 MDRPDSYNSFLICVANDVSESETRAKFESLFRITYDKOTTFQYKSFKRVINFSN 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 4 VDLQDLP--SATIACHLDPRVFDGLCRAKFESLFRITYDKOTTFQYKSFKRVINFSN 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 61 PLSAADARLRLHKTETFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPPVGMKQV 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 PLSAADARLRLHKTETFLGKEMKLYFAQTLHGSSHLAPPNDKQFLISPPASPPVGMKQV 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 121 EDATPVINVDLLVAISKLGPEKYTELHAATDTPPSVVHVHCSDQENEEBEMERMRKP 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 121 EDATPVINVDLLVAISKLGPEKYTELHAATDTPPSVVHVHCSDQENEEBEMERMRKP 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 181 KPKIOTRRPEYTPHLS 198
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 181 KPKIOTRRPEYTPHLS 198
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 12

```
US-09-782-953-16
; Sequence 16, Application US/09782953
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)
; FILE REFERENCE: UTSD:674P21
; CURRENT APPLICATION NUMBER: US/09/782,953
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-16

Query Match      83.9%; Score 875; DB 21; Length 197;
Best Local Similarity 86.4%; Pred. No. 2.3e-78;
Matches 171; Conservative 8; Mismatches 15; Indels 4; Gaps 2;

QY 1 MDRPDSYFNSSLIACVANDVSESETRAKFESLRTYDKDTTFQYFKSFKVRINFNSN 60
DB 4 VDLQDLP---SATIACHLDPRVFVGLCRAKFESLRTYDKDTTFQYFKSFKVRINFNSN 60

QY 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHGSSHLAPPNDPKQFLISPPASPPVGWKQV 120
DB 61 PLSAADARLQLHKTFLGKEMKLYFAQTLHGSSHLAPPNDPKQFLISPPASPPVGWKQV 120

QY 121 EDATPVINYDLLVAISKLGPGKVELHAATDTPPSVVHVHVCSDQENEESEMERMRKP 180
DB 121 EDATPVINYDLLVAISKLGPGKVELHAATDTPPSVVHVHVCSDQENEESEMERMRKP 179

QY 181 KPKIOTRRPEYTPHLS 198
DB 180 KPKIOTRRPEYTPHLS 197

RESULT 13
US-10-104-047-2892
; Sequence 2892, Application US/10104047
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cdna
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2892
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-2892

Query Match      83.9%; Score 875; DB 25; Length 252;
Best Local Similarity 86.4%; Pred. No. 3.2e-78;
Matches 171; Conservative 8; Mismatches 15; Indels 4; Gaps 2;

QY 1 MDRPDSYFNSSLIACVANDVSESETRAKFESLRTYDKDTTFQYFKSFKVRINFNSN 60
DB 59 VDLQDLP---SATIACHLDPRVFVGLCRAKFESLRTYDKDTTFQYFKSFKVRINFNSN 115

QY 61 PLSAADARLRLHKTFLGKEMKLYFAQTLHGSSHLAPPNDPKQFLISPPASPPVGWKQV 120
DB 116 PLSAADARLQLHKTFLGKEMKLYFAQTLHGSSHLAPPNDPKQFLISPPASPPVGWKQV 175

QY 121 EDATPVINYDLLVAISKLGPGKVELHAATDTPPSVVHVHVCSDQENEESEMERMRKP 180
DB 121 EDATPVINYDLLVAISKLGPGKVELHAATDTPPSVVHVHVCSDQENEESEMERMRKP 180

US-09-782-953-7
; Sequence 7, Application US/09575580B
; GENERAL INFORMATION:
; APPLICANT: McKeon, F.
; APPLICANT: Kayako, K.
; APPLICANT: Ryeom, S.
; TITLE OF INVENTION: CALCIPRESSINS: ENDOGENOUS INHIBITORS OF CALCINEURIN
; TITLE OF INVENTION: USES AND REAGENTS RELATED THERETO
; FILE REFERENCE: HMV-048.01
; CURRENT APPLICATION NUMBER: US/09/575,580B
; CURRENT FILING DATE: 2000-05-22
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-575-580B-7

Query Match      80.9%; Score 843.5; DB 19; Length 170;
Best Local Similarity 95.8%; Pred. No. 2.6e-75;
Matches 161; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

QY 30 AKFESLFRDYDKDTTFQYFKSFKVRINFNSNPLSADARLRLHKTFLGKEMKLYPAQTL 89
DB 4 AKFESLFRDYDKDTTFQYFKSFKVRINFNSNPLSADARLRLHKTFLGKEMKLYPAQTL 63

QY 90 HIGSSHLAPPNDPKQFLISPPASPPVGWKQVEDATPVINYDLYAISKLGPGEKVELHAA 149
DB 64 HIGSSHLAPPNDPKQFLISPPASPPVGWKQVEDATPVINYDLYAISKLGPGEKVELHAA 123

QY 150 TDTPSVVVHVHVCSDQENEESEMERMRKPXPKIOTRRPEYTPHLS 197
DB 124 TDTPSVVVHVHVCSDQENEESEMERMRKPXPKIOTRRPEYTPHLS 170

RESULT 15
US-09-782-953-9
; Sequence 9, Application US/09782953
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)
; FILE REFERENCE: UTSD:674P21
; CURRENT APPLICATION NUMBER: US/09/782,953
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-782-953-9

Query Match      60.6%; Score 632.5; DB 21; Length 197;
Best Local Similarity 64.4%; Pred. No. 3.4e-54;
Matches 121; Conservative 24; Mismatches 34; Indels 9; Gaps 2;

QY 11 SSLIACVANDVSESETRAKFESLRTYDKDTTFQYFKSFKVRINFNSNPLSADARL 70
DB 11 STLVAACVVDVEVFTNQEVKEXPEGLFRTYDECVTFQLFKSFRRVRINFNSHPKSAARARIE 70
```

QY 71 LKTEFLGKEMKLYFAQTLHIGSS---HLAPPNDKQFLISPPASPPVGMKQVEDATPV 126
Db 71 LHETQFRGKKLKLYPAQVCTPETDGDKLHLAPPQPAKQFLISPPSSPPVGMKPISDATPV 130
QY 127 INYDLLIAISKLGPGKYEHLAATDPTPSVVVHVCESDOENESEEMERMKRPPKIIQ 186
Db 131 LNYDLLYAVAKLGPGKYEHLAAGTESVHVHVCDSDEEDPK----TSPKIIQ 185
QY 187 TRREYTP 194
Db 186 TRRGGLPP 193

Search completed: December 11, 2002, 11:49:34
Job time : 137.563 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 14, 2002, 20:00:06 ; Search time 30.0506 Seconds
(without alignments)
2020.659 Million cell updates/sec

Title: US-09-782-953-7

Perfect score: 1043

Sequence: 1 MDRFDSYNFSSLIACVAND.....RPAKLIQTRRPEYTPHLS 198

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=framet_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US0782953/runat 11122002 114430 17377/app_query.fasta_1.1173
-DB=issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPECL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTPMT=ptc -NORM=ext -HAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0782953 @CEN_1_1_29 @runat 11122002 114430 17377 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MWAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTCUS COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	844.5	81.0	2174	2	US-08-665-040-1
2	94	9.0	1820	5	PCN-US94-09752-1
3	81	7.8	3138	3	US-09-234-332-5
4	81	7.8	3728	1	US-08-111-939-1
5	81	7.8	5467	2	US-07-745-206A-12
6	81	7.8	5467	2	US-08-311-363-12
7	81	7.8	7175	1	US-08-455-543A-8
8	81	7.8	7175	2	US-08-193-078B-8
9	81	7.8	7175	2	US-08-223-305C-8
10	81	7.8	7175	2	US-08-149-097D-8
11	81	7.8	7175	2	US-08-949-386-8
12	81	7.8	7175	3	US-08-450-562-8
					Sequence 1, Appli
					Sequence 1, Appli
					Sequence 5, Appli
					Sequence 12, Appli
					Sequence 12, Appli
					Sequence 8, Appli
					Sequence 8, Appli
					Sequence 8, Appli
					Sequence 8, Appli

13	81	7.8	7175	4	US-08-984-709A-8	Sequence 8, Appli
14	81	7.8	7175	4	US-08-450-272-8	Sequence 8, Appli
15	81	7.8	7177	4	US-09-268-163-7	Sequence 7, Appli
16	81	7.8	7362	1	US-08-455-543A-7	Sequence 7, Appli
17	81	7.8	7362	2	US-08-193-078B-7	Sequence 7, Appli
18	81	7.8	7362	2	US-08-223-305C-7	Sequence 7, Appli
19	81	7.8	7362	2	US-08-149-097D-7	Sequence 7, Appli
20	81	7.8	7362	3	US-08-949-386-7	Sequence 7, Appli
21	81	7.8	7362	3	US-08-450-562-7	Sequence 7, Appli
22	81	7.8	7362	4	US-08-984-709A-7	Sequence 7, Appli
23	81	7.8	7362	4	US-08-450-272-7	Sequence 7, Appli
24	81	7.8	7364	4	US-09-268-163-5	Sequence 5, Appli
25	81	7.8	7376	4	US-09-268-163-3	Sequence 3, Appli
26	80.5	7.7	6232	4	US-08-456-200B-11	Sequence 11, Appli
27	80	7.7	1113	1	US-08-231-342-5	Sequence 5, Appli
28	80	7.7	3093	1	US-08-252-966B-19	Sequence 19, Appli
29	79	7.6	1113	1	US-08-231-342-7	Sequence 7, Appli
30	79	7.6	36651	4	US-09-738-894A-3	Sequence 3, Appli
31	78.5	7.5	1644	4	US-09-877-730-25	Sequence 25, Appli
32	78.5	7.5	1875	4	US-09-877-730-23	Sequence 23, Appli
33	78.5	7.5	1887	4	US-09-877-730-29	Sequence 29, Appli
34	78.5	7.5	2139	4	US-09-877-730-21	Sequence 21, Appli
35	78.5	7.5	2382	4	US-09-877-730-27	Sequence 27, Appli
36	78.5	7.5	2481	4	US-09-877-730-15	Sequence 15, Appli
37	78.5	7.5	2715	4	US-09-877-730-5	Sequence 5, Appli
38	78.5	7.5	2724	4	US-09-877-730-19	Sequence 19, Appli
39	78.5	7.5	2958	4	US-09-877-730-9	Sequence 9, Appli
40	78.5	7.5	2976	4	US-09-877-730-11	Sequence 11, Appli
41	78.5	7.5	3210	4	US-09-877-730-1	Sequence 1, Appli
42	78.5	7.5	3219	4	US-09-877-730-17	Sequence 17, Appli
43	78.5	7.5	3453	4	US-09-877-730-7	Sequence 7, Appli
44	78.5	7.5	3874	4	US-09-877-730-31	Sequence 31, Appli
45	78.5	7.5	6171	1	US-08-459-568-1	Sequence 1, Appli

ALIGNMENTS

US-08-665-040-1
; Sequence 1, Application US/08665040
; Patent No. 5869118
; GENERAL INFORMATION:
; APPLICANT: ESTIVILL PALLEJA, XAVIER
; APPLICANT: FUENTES, JUAN JOSE
; APPLICANT: PRITCHARD, MELANIE
; TITLE OF INVENTION: A NEW GENE SEQUENCE OF THE
; TITLE OF INVENTION: DOWN SYNDROM CRITICAL REGION OF HUMAN CHROMOSOME 21,
; TITLE OF INVENTION: IDENTIFIED BY A NEW "ALU-SPLICING PCR" TECHNIQUE,
; TITLE OF INVENTION: CODING FOR A PROLINE-RICH PROTEIN (DSCR1) HIGHLY
; TITLE OF INVENTION: EXPRESSED IN FOETAL BRAIN AND IN HEART AND METHOD
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LADAS & PARRY
; STREET: 26 WEST 61ST STREET
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10023
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORD PERFECT 5.1 FOR DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/665,040
; FILING DATE: JUNE 7, 1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: ES P9501140
; FILING DATE: JUNE 7, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: JANET I. CORD

REGISTRATION NUMBER: 33,778
 REFERENCE/DOCKET NUMBER: U010815-9
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 708-1800
 TELEFAX: (212) 246-9959
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2174 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: circular
 MOLECULE TYPE: cDNA for mRNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE: human
 ORGANISM: Homo sapiens
 DEVELOPMENTAL STAGE: foetal
 TISSUE TYPE: Brain
 IMMEDIATE SOURCE: gene library of cDNA
 LIBRARY: gene library of cDNA from foetal
 CLONE: BC-17.8-1 and BC-17.8-2
 POSITION IN GENOME:
 CHROMOSOME/SEGMENT: chromosome 21/YAC 72H9
 MAP POSITION: 21q22.1-q22.2
 FEATURE:
 NAME/KEY: cDNA for mRNA, BC-17.8
 LOCATION: 1..2174
 FEATURE:
 NAME/KEY: untranslated 5'
 LOCATION: 1..48
 FEATURE:
 NAME/KEY: coding sequence
 LOCATION: 49..560
 OTHER INFORMATION: Down Syndrome critical
 region 1 (DSCR1)
 FEATURE:
 NAME/KEY: DSCR1
 LOCATION: 1..171 PEPTIDES
 IDENTIFICATION METHOD: translation of the
 OTHER INFORMATION: DSCR1 sequence.
 OTHER INFORMATION: - deduced protein
 OTHER INFORMATION: - proline-rich protein domains
 OTHER INFORMATION: - glutamic acid-rich protein domains
 OTHER INFORMATION: - leucine/phenylalanine-rich protein domains
 FEATURE:
 NAME/KEY: untranslated 3'
 LOCATION: 564..2174
 FEATURE:
 NAME/KEY: 2 poly (A)
 LOCATION: 1541..1546 AND 2132..2137
 US-08-665-040-1
 Alignment Scores:
 Pred. No.: 1.95e-98 Length: 2174
 Score: 844.50 Matches: 161
 Percent Similarity: 97.63% Conservative: 4
 Best Local Similarity: 95.27% Mismatches: 3
 Query Match: 80.97% Indels: 1
 Gaps: 2
 DB: 1
 US-09-782-953-7 (1-198) x US-08-665-040-1 (1-2174)
 QY 30 AlAlaYsPheGluSeRLeuPheArgThrTYrAspLYsApThrThrPheGlnTYrPheLYs 49
 DB 58 GCCAAATTGAGTCCCTTATGAGACGTATGACAGACATCACTTTCAGTATTAAAG 117
 QY 50 SerPheLYsArgValAlaGllLeaenPheSerRanProLeuSerAlaAlaAaPAlaArgLeu 69
 DB 118 AGCTTCAACAGAGTCACAAATTAATCTCAGCAACCCCTTCCGAGAGATGCCAGGCTC 177
 QY 70 ArgLeuHIsLYsErThrGluPheLeuGlyLeuGluMetLYsLeuTYrPheAlaGlnThrLeu 89
 DB 178 CAGCTGCAATAGACTGAGTTCTCGGAAAGAAATGAAGTTATTTTGTCTCAGACCTTA 237

QY 90 HisIleGlySerSerHisLeuAlaProProAsnProAspLYsGlnPheLeuIleSerPro 109
 DB 238 CACATAGAGAGCTCACACCTGGGCTCCGCCAAATCCAGACAGAGATTTCTGATCTCCCT 297
 QY 110 ProIleSerProProValGlyTYrPlyGlnValAlaGAspAlaThrProValIleAsnTYr 129
 DB 298 CCGCCTCTCCGACGATGGGATGGAAACATCGAAGATCGACCCCAAGCTAAACTAT 357
 QY 130 AspLeuLeuTYrAlaIleSerLYsLeuGlyProGlyGluLYsTYrGluLeuHisAlaAla 149
 DB 358 GATCTTTATATGATCATCTCCAAAGCTGGGGCCAGGGGAAAGTATGATTCAGACGAGCG 417
 QY 150 ThrAspProThrProSerValValHisValCysGluSerAspGlnGluAanGluGlu 169
 DB 418 ACTACACACACTCCACGCTGGTGTCTCATGTATGTAGAGTATCAAGH---AAGGAG 474
 QY 170 GluGluGluGluMetGluAArgMetLYsArgProLYsProLYsIleIleGlnThrArgArg 189
 DB 475 GAAGAAAGGAAATGAAAGAAATAGAGAGACCTTAAGCAAAATTTATCCAGACGAGAGG 534
 QY 190 ProGluTYrThrProIleHisLeuSer 198
 DB 535 CCGAGTACACGCCGATCCACTCAGC 561
 RESULT 2
 PCT-US94-09752-1
 Sequence 1, Application PC/TUS9409752
 GENERAL INFORMATION:
 APPLICANT: David S. Strayer and Avinash Chander
 TITLE OF INVENTION: Compositions and Methods for
 Targeting Cells and Modulating Pulmonary Surfactant Secretion
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Jane Maasey Licata, Esq.
 STREET: 210 Lake Drive East, Suite 201
 CITY: Cherry Hill
 STATE: NJ
 COUNTRY: USA
 ZIP: 08002
 COMPUTER READABLE FORM:
 MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB
 MEDIUM TYPE: STORAGE
 COMPUTER: IBM PS/2
 OPERATING SYSTEM: PC-DOS
 SOFTWARE: WORDPERFECT 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US94/09752
 FILING DATE: Herewith
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/176,218
 FILING DATE: December 30, 1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/114,951
 FILING DATE: August 31, 1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Jane Maasey Licata
 REGISTRATION NUMBER: 32,257
 REFERENCE/DOCKET NUMBER: JEFF-0042
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (609) 779-2400
 TELEFAX: (609) 779-8488
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1820
 TYPE: nucleic acid
 STRANDEDNESS: single stranded
 TOPOLOGY: linear
 PCT-US94-09752-1
 Alignment Scores:
 Pred. No.: 0.0363 Length: 1820
 Score: 94.00 Matches: 49

Alignment Scores:

REFERENCE/DOCKET NUMBER:

REGISTRATION NUMBER: 33,694
REFERENCE/DOCKET NUMBER: 02481.1321-00000
TELECOMMUNICATION INFORMATION:

```

TELEPHONE: 202-408-4000
TELEFAX: 202-408-4000
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3728 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: Mus musculus
STRAIN: osteoblastic cell line MC3T3E1
FEATURE:
NAME/KEY: CDS
LOCATION: 69..3452
US-08-111-939-1

Alignment Scores:
Pred. No.: 5.13 Length: 3728
Score: 81.00 Matches: 33
Percent Similarity: 38.99% Conservative: 29
Best Local Similarity: 20.75% Mismatches: 37
Query Match: 7.77% Indels: 60
DB: 1 Gaps: 8

US-09-782-953-7 (1-198) x US-08-111-939-1 (1-3728)

QY 50 SerPheLyArgValArg-----1le 56
DB 2997 AACTGGAAGCGCATTCGGAGATCTTGCTATGACGGAAACCGTCCCTTCGAGATT 3056
QY 57 AenPheSerAnPheLeuSerAlaAlaPheAlaArgLeuArgLeuHisLeuThrGluPhe 76
DB 3057 GACCCCTCAGACACCATGACGCCCCCAGCGCGCATGAGAGCGCGCTACAGTAC 3116
QY 77 -----LeuGlyysGluMetLysLeuTyPheAlaGlnThrLeuHisLeuGlySer 93
DB 3117 CGGCTCCGATGAGGAAACAGATGCAACTG-----CGTGGCTCAATCTTACCGCA 3167
QY 94 SerHisLeuAlaProPheAnPheProAlaPheGlnPheLeuLeuSerProProAlaSerPro 113
DB 3168 GGCCCTGCCACAAAGCCCACTCT-----GCCCTTATGCTCTCCCTTCCCT 3215
QY 114 -----ProValGlyTyPlys 118
DB 3216 ACACGACCATTAACCTTGAGGCCCTGGAAATTCTACCACTACCACTGAGCGCTGGAG 3275
QY 119 -----GlnValGluAraAlaThrProValIleAsnTyTrsPheLeuLeuTyAlaIleSer 136
DB 3276 GAGTCAGAGACTGACCTTACAGAGTAGTGACA-----3311
QY 137 LysLeuLeuTyProGlyGluLysTyGluLeuHisLeuAlaIleThrAspProThrProSerVal 156
DB 3312 -----GAGTTTGAACAGAGATAGGACTGAC-----3338
QY 157 ValValHisValCysGluSerAspGlnGluAsnGluGluGluGluGluMetGlu 175
DB 3339 ---CTAGAGGTGAAAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3392

RESULT 5
US-07-745-206A-12
Sequence 12, Application US/07745206A
Patent No. 5429921
GENERAL INFORMATION:
APPLICANT: Harpold, Michael
APPLICANT: Ellis, Steven
APPLICANT: Williams, Mark
APPLICANT: McCue, Ann
APPLICANT: Feldman, Daniel
TITLE OF INVENTION: Human Calcium Channel Compositions and
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:

```

```

ADDRESSEE: Fitch, Even, Tabin & Flannery
STREET: 135 S. LaSalle
CITY: Chicago
STATE: Illinois
COUNTRY: U.S.A.
ZIP: 60603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/745,206A
FILING DATE: 19910815
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Feder, Scott B
REFERENCE/DOCKET NUMBER: 51504
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-372-7842
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 5467 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: join(144..3164, 3168..3245, 3249..3386, 3390
LOCATION: ..3392, 3396..3489, 3495..3539, 3543..3581, 3585
LOCATION: ..3587, 3591..3626, 3630..3689, 3693..3737, 3744
LOCATION: ..3746, 3750..4823, 4827..4841, 4845..5006, 5010
LOCATION: ..5096, 5100..5306, 5310..5366, 5370..5465)
US-07-745-206A-12

Alignment Scores:
Pred. No.: 9.26 Length: 5467
Score: 81.00 Matches: 50
Percent Similarity: 33.77% Conservative: 27
Best Local Similarity: 21.93% Mismatches: 80
Query Match: 7.77% Indels: 71
DB: 1 Gaps: 9

US-09-782-953-7 (1-198) x US-07-745-206A-12 (1-5467)

QY 7 SerTyraenPheSerSerLeuIleAlaCysValAlaAenAraPheValPheSerGluSer 26
DB 2232 AACTACACTGCTGTAAGATGCTTCTTGCGCATCGCTGTGCAACCTGGCCACGCCAA 2291
QY 27 GluThrArgAlaLysPheGluSerLeuPheArgThrTyraPheLysAspThrThrPheGln 46
DB 2292 GAGCTGACCAAGATGAAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2351
QY 47 TyrPheLysSerPheLysArgValArgIleAsnPheSerAnProLeuSerAlaAlaAsp 66
DB 2352 AAGCCCAAGAGAGTGTGAGTGC-----AAGCCCATGTGTGCGCGCAAC 2396
QY 66 -----66
DB 2397 ATTCATGCGCCGACGAGAGAGAACTCGGCCAAGGCGCGCTGGTGGAGAGAGCGG 2456
QY 67 ---AlaArgLeuArgLeuHis-----LysThrGluPheLeuGlyLysGluMet 81
DB 2457 GCCAGCAGCTACGGCTGAGAACTCGGGGCGCAGCTGAGAGCGCTGTACAGCGAGATG 2516
QY 82 -----LysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAla 97
DB 2517 GACCCGAGAGAGCGGTGGCTTGGCACTACGCGCCACCTG-----2558
QY 98 ProProAnPheAspLysGlnPheLeuIleSerProPro-----110
DB 2559 -----CGGCCGACATGAAGAGCAGACCTGACCGCGCTGGTGTGAGAGCTGGAGCGCC 2612

```


TITLE OF INVENTION: METHODS
 NUMBER OF SEQUENCES: 57
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Brown, Martin, Haller & McClain
 STREET: 1660 Union Street
 CITY: San Diego
 STATE: California
 COUNTRY: USA
 ZIP: 92101-2926
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FASTSEQ Version 1.5
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/455,543A
 FILING DATE: May 31, 1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/223,305
 FILING DATE: April 4, 1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/868,354
 FILING DATE: April 10, 1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/745,206
 FILING DATE: 15-AUG-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/620,250
 FILING DATE: 30-NOV-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/482,384
 FILING DATE: 20-FEB-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/603,751
 FILING DATE: 04-APR-1989
 APPLICATION NUMBER: WO PCT/US89/01408
 FILING DATE: 04-APR-1989
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/1176,899
 FILING DATE: 04-APR-1988
 ATTORNEY/AGENT INFORMATION:
 NAME: Seidman, Stephanie L.
 REGISTRATION NUMBER: 33,779
 REFERENCE/DOCKET NUMBER: 6362-52517
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619)238-0999
 TELEFAX: (619)238-0062
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 7175 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 144..6857
 NAME/KEY: 5'UTR
 LOCATION: 1..143
 FEATURE:
 NAME/KEY: 3'UTR
 LOCATION: 6855..7175
 US-08-455-543A-8

Alignment Scores:
 Pred. No.: 14.1
 Score: 81.00
 Percent Similarity: 33.79%
 Best Local Similarity: 21.46%
 Query Match: 7.77%
 DB: 1

Length: 7175
 Matches: 47
 Conservative: 27
 Mismatches: 75
 Indels: 70
 Gaps: 9

US-09-782-953-7 (1-198) x US-08-455-543A-8 (1-7175)
 QY 7 SerTyrAsnPhseSerLeuIleAlaCysValAlaAsnAspAspValPheSerGluSer 26
 Db 2232 AACCTACACTCTGTGTGAATGCTTTCTTGGCCATCGGCTGTGACAACTGGCCAAAGCCCA 2291
 QY 27 GluThrArgAlaLysPheGluSerLeuPheArgThrTyrAspLysAspThrPheGln 46
 Db 2292 GAGCTGACCAAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 2351
 QY 47 TyrPheLysSerPheLysArgValArgGlyLeuAsnPhseSerAspProLeuSerAlaAlaAsp 66
 Db 2352 AAGGCCAAGAGTGGCTGAATC-----AGCCCATGTCTCTCCGGGAAC 2396
 QY 66 ----- 66
 Db 2397 ATCTCCATCGCCGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 2456
 QY 67 ---AlaArgLeuArgLeuHis-----LysThrGluPheLeuGlyLysGluMet 81
 Db 2457 GCCAGCCAGCTACGCTGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 2516
 QY 82 -----LysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAla 97
 Db 2517 GACCCCGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 2558
 QY 98 ProProAsnProAspLysGlnPheLeuIleSerProPro----- 110
 Db 2559 -----CGGCCGACATGAAAGACGACGACGACGACGACGACGACGACGACGACGACGACG 2612
 QY 111 -----AlaSerProProValGlyTyrLysGlnValGlnAspAlaThrProValIleAsn 128
 Db 2613 GACG 2663
 QY 129 TyrAspLeuLeuTyrAlaIleSerLysLeuGlyProGlyLysTyrGluLysIleAla 148
 Db 2664 -----GCCCGGAGGGGTGACCTCGCGGACGACGACGACGACGACGACGACGACGACG 2708
 QY 149 AlaThrAspProThrProSerValValValHisValLysGluSerAspGlnGlu----- 166
 Db 2709 GACAAGACAAAGACCCCGCGCGCG-----GGGACACGAGACCGACGACGACGACGACGACG 2750
 QY 167 -----AanglGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 182
 Db 2751 GAGGCCCGGAAGCGGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 2807

RESULT 8
 US-08-193-078B-8
 Sequence 8, Application US/08193078B
 Patent No. 5846757
 GENERAL INFORMATION:
 APPLICANT: Harpold, Michael
 APPLICANT: Ellis, Steven
 APPLICANT: Williams, Mark
 APPLICANT: Feldman, Daniel
 APPLICANT: McCue, Ann
 APPLICANT: Brenner, Robert
 TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
 NUMBER OF SEQUENCES: 29
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: BROWN, MARTIN, HALLER & MCCLAIN
 STREET: 1660 UNION STREET
 CITY: SAN DIEGO
 STATE: CA
 COUNTRY: USA
 ZIP: 92101
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/193,078B
 ; FILING DATE: 07-FEB-1994
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/868,354
 ; FILING DATE: 10-APR-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/745,206
 ; FILING DATE: 15-AUG-1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Seidman, Stephanie L.
 ; REGISTRATION NUMBER: 33,779
 ; REFERENCE/DOCKET NUMBER: 6362-53607
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 619-238-0999
 ; TELEFAX: 619-238-0062
 ; INFORMATION FOR SEQ ID NO: 8:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 7175 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 144..6857
 ; FEATURE:
 ; NAME/KEY: 5'UTR
 ; LOCATION: 1..143
 ; FEATURE:
 ; NAME/KEY: 3'UTR
 ; LOCATION: 6855..7175
 ; US-08-193-078B-8

Alignment Scores:
 Pred. NO.: 14.1 Length: 7175
 Score: 81.00 Matches: 47
 Percent Similarity: 33.79% Conserved: 27
 Best Local Similarity: 21.46% Mismatches: 75
 Query Match: 7.77% Indels: 70
 DB: 2 Gaps: 9

US-09-782-953-7 (1-198) x US-08-193-078B-8 (1-7175)

Qy 7 serTyrAsnProPheSerLeuLeuAlaCysValAlaAsnAspValPheSerGluSer 26
 Db 2332 AACTACACTGCTGATGATGCTTTCTGCCATCTGCTGGACAACTGGCCAAAGCCCAA 2291
 Qy 27 GluThrArgAlaLysPheGluSerLeuPheArgThrTyrAspLysAspThrThrPheGln 46
 Db 2292 GAGCTGACCAAGGATGAAGAGAGATGGAAGAGCAGCCCAATCAGAGCTTGTCTGCAA 2351
 Qy 47 TyrPheLysSerPheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAsp 66
 Db 2352 AAGGCCAAAGAGTGGCTGAATC-----AGCCCCATGCTGCCGCGAAC 2396
 Qy 66 ----- 66
 Db 2397 ATCTCCATCCGCCAGGACGAGAACTCGGCCAAGCGCGCTCGGTGGGAGCAGCGG 2456
 Qy 67 -----AlaArgLeuArgLeuHis-----LysThrGluPheLeuGlyLysGluMet 81
 Db 2457 GCCAGCCAGCTACGGCTGCAGAACCTCGCGGCGAGCTGCGGCGCTGTACAGCGAGATG 2516
 Qy 82 -----LysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAla 97
 Db 2517 GACCCGAGAGCGGTGGCTTTCGCCACTACGCGCACCTG----- 2558
 Qy 98 ProProAsnProAspLysGlnPheLeuIleSerProPro----- 110
 Db 2559 -----CGGCCCCACATGAGAGCGACCTGGACCGCCCGCTGGTGTGTAGCTGGCGCGC 2612

Qy 111 -----AlaSerProProValGlyTyrLysGlnValGluAspAlaThrProValIleAsn 128
 Db 2613 GACGGCGCGGGGGCGCGCTGGGAGGCAAGCCGACCTGAGGCTCGGAG----- 2663
 Qy 129 TyrAspLeuLeuTyrAlaIleSerLysLeuGlyProGlyGluLysTyrGluLeuHisAla 148
 Db 2664 -----GCCCCGAGGGGCTGACCCCTCCGCGCAGGACCCACCGGACCGC 2708
 Qy 149 AlaThrAspProThrProSerValValHisValCysGluSerAspGlnGlu----- 166
 Db 2709 GACAAGGACAAGACCCCGCGCG-----GGGACCCAGGACCGACCA 2750
 Qy 167 -----AsnGluGluGluGluGluMetGluArgMetLysArgProLysPro 182
 Db 2751 GAGGCCCGAAGCGGAGAGCGGGAGCCCGTGGCCGCGGAGGCGCGCGCGCG 2807
 RESULT 9
 US-08-223-305C-8
 ; Sequence 8, Application US/08223305C
 ; Patent No. 5851824
 ; GENERAL INFORMATION:
 ; APPLICANT: Harpold, Michael
 ; APPLICANT: Ellis, Steven
 ; APPLICANT: Williams, Mark
 ; APPLICANT: Feldman, Daniel
 ; APPLICANT: McCue, Ann
 ; APPLICANT: Brenner, Robert
 ; TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
 ; TITLE OF INVENTION: METHODS
 ; NUMBER OF SEQUENCES: 57
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Brown, Martin, Haller & McClain
 ; STREET: 1660 Union Street
 ; CITY: San Diego
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 92101-2926
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSEQ Version 1.5
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/223,305C
 ; FILING DATE: April 4, 1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 07/868,354
 ; FILING DATE: April 10, 1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/745,206
 ; FILING DATE: 15-AUG-1991
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/620,250
 ; FILING DATE: 30-NOV-1990
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/482,384
 ; FILING DATE: 20-FEB-1990
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/603,751
 ; FILING DATE: 04-APR-1989
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: WO PCT/US89/01408
 ; FILING DATE: 04-APR-1989
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/176,899
 ; FILING DATE: 04-APR-1988
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Seidman, Stephanie L.
 ; REGISTRATION NUMBER: 33,779
 ; REFERENCE/DOCKET NUMBER: 52516 (P519739)
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (619)238-0999
 ; TELEFAX: (619)238-0062

INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 7175 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 144..6857
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..143
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 6855..7175
US-08-223-305C-8
Alignment Scores:
Pred. No.: 14.1 Length: 7175
Score: 81.00 Matches: 47
Percent Similarity: 33.79% Conservative: 27
Best Local Similarity: 21.46% Mismatches: 75
Query Match: 7.77% Indels: 70
Gaps: 9
DB:
US-09-782-953-7 (1-198) x US-08-223-305C-8 (1-7175)
QY 7 SerTyrAsnPheserSerLeuIalaCyValAlaAsnAspValPheserGluSer 26
DB 2232 AACTACACTGTGTAATCTTCTGGCCATGCGACAACTGCGCAAGCCCAA 2291
QY 27 GluThrArgAlaYserGluSerLeuPheArgThrTyrAspIysAspThrThrPheGln 46
DB 2292 GAACGTGACCAAGAGATGAAGAGAGATGAAGAAAGCAAGCCATAGAACTGCTGCA 2351
QY 47 TyrPheIysSerPheIysValArgIleAsnPheserAsnProIysSerAlaIaAsp 66
DB 2352 AAGCCCAAAAGTGGCTGAATC-----AGCCCAATGCTGCGCGCAAC 2396
QY 66 ----- 66
DB 2397 ATCTCATCCGCCCGACAGCAAGAACTCGAGCGCGCTGCTGGAGCAGCGG 2456
QY 67 ---AlaArgLeuArgLeuHis-----LysThrGluPheLeuGlyLysGluMet 81
DB 2457 GCCAGCGACGTACGCGCTGCAGAACTCGCGGCGCAGCTGCGAGCGCTGACAGAGATG 2516
QY 82 -----LysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAla 97
DB 2517 GACCCCGAGAGCGGCGTGGCTTCGACCTACGCGCACCTG----- 2558
QY 98 ProProAsnProAspLysGlnPheLeuIleSerProPro----- 110
DB 2559 -----CGGCGCGACATGAAGACCGACCTGAGCGCGCGCTGCTGCTGAGCTGGCGCG 2612
QY 111 -----AlaSerProProValGlyTyrPheGlnValGluAspAlaThrProValIleAsn 128
DB 2613 GACGCGCGCGCGCGCGCTGCGAGCGCAAGCCGACCTGAGCGTGGAG----- 2663
QY 129 TyrAspLeuLeuTyrAlaIleSerIysLeuGlyProGlyGluLysTyrGluLeuHisAla 148
DB 2664 -----GCCCGCGAGGCGCTGACCTCGCGCACGACACACCGCGCACCGC 2708
QY 149 AlaThrAspProThrProSerValValValHisValCysGluSerAspGlnGlu----- 166
DB 2709 GACAAGACCAAGACCCCGCGCG-----GGGACCGAGACCGAGCA 2750
QY 167 -----AsnGluGluGluGluGluGluMetGluArgMetLysArgProLysPro 182
DB 2751 GAGCGCCGAGAGCGGAGAGCGGCGGAGCGCGTCCCGAGAGAGCGCGCGCGCG 2807
RESULT 10

US-08-149-097D-8
Sequence 8, Application US/08149097D
Patent No. 5874236
GENERAL INFORMATION:
APPLICANT: Harpold, Michael
APPLICANT: Ellis, Steven
APPLICANT: Williams, Mark
APPLICANT: Feldman, Daniel
APPLICANT: McCue, Ann
APPLICANT: Brenner, Robert
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
TITLE OF INVENTION: METHODS
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brown, Martin, Haller & McClain
STREET: 1660 Union Street
City: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101-2926
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/149,097D
FILING DATE: 05-NOV-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/105,536
FILING DATE: 11-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US92/06903
FILING DATE: 14-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/914,231
FILING DATE: 13-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/868,354
FILING DATE: 10-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/745,206
FILING DATE: 15-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/620,250
FILING DATE: 30-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/482,384
FILING DATE: 20-FEB-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/603,751
FILING DATE: 04-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US89/01408
FILING DATE: 04-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/176,899
FILING DATE: 04-APR-1988
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 6362-55038
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 238-0999
TELEFAX: (619) 238-0062
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 7175 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)

```

; FEATURE: CDS
; NAME/KEY: 144..6857
; LOCATION: 144..6857
; FEATURE: 5'UTR
; NAME/KEY: 1..143
; LOCATION: 1..143
; FEATURE: 3'UTR
; NAME/KEY: 6855..7175
; LOCATION: 6855..7175
US-08-149-097D-8

Alignment Scores:
Pred. No.: 14.1 Length: 7175
Score: 81.00 Matches: 47
Percent Similarity: 33.79% Conservative: 27
Best Local Similarity: 21.46% Mismatches: 75
Query Match: 7.77% Indels: 70
DB: 2 Gaps: 9

US-09-782-953-7 (1-198) x US-08-149-097D-8 (1-7175)
Qy 7 SerTyrAsnPheSerLeuLeuAlaCysValAlaAsnAspValPheSerGluSer 26
Db 2232 AACTACACTGCTGGAATGTCTTTCTGCCATCGCTGGACAACCTGGCCACGCCCAA 2291
Qy 27 GluThrArgAlaLysPheGluSerLeuPheArgThrTyrAspLysAspThrPheGln 46
Db 2292 GAGCTGACCAAGGATGAAGAGGAGATGAAGAGCAGCCATCAGAACCTTGCTCTGCAA 2351
Qy 47 TyrPheLysSerPheLysArgValArgLeuAlaAsnPheSerAsnProLeuSerAlaAlaAsp 66
Db 2352 AAGCCAAAGAGTGGCTGAAGTC-----AGCCCCATGTCTGCCCGGAC 2396
Qy 66 -----66
Db 2397 ATCTCCATCGCCGACGACGACAGACACTCGGCCAAGCGCGCTCGGTGGAGACGCG 2456
Qy 67 ---AlaArgLeuArgLeuHis-----LysThrGluPheLeuGlyLysGluMet 81
Db 2457 GCCAGCCAGCTACGCTGCAGAACCTCGGGCCAGCTGCGAGGCGCTGTACAGCGAGATG 2516
Qy 82 -----LysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAla 97
Db 2517 GACCCCGAGGAGCGCTGCGCTTCCGCACTACGCGCCACCTG-----2558
Qy 98 ProProAsnProAspLysGlnPheLeuLeuSerProPro-----110
Db 2559 -----CGGCCCGACATGAAGACGACCTCGACCGCGCGCTGCTGGAGTGGCGCGC 2612
Qy 111 -----AlaSerProProValGlyTyrLysGlnValGluAspAlaThrProValIleAsn 128
Db 2613 GACGGCGCGCGCGCGCGCGTGGAGGCAAGCCGACCTGAGCTCGGAG-----2663
Qy 129 TyrAspLeuLeuTyrAlaIleSerLysLeuGlyProGlyGluLysTyrGluLeuHisAla 148
Db 2664 -----GCCCGCGAGGCGCTCGACCTCGCGCGAGCGCACCCACCGCACCGC 2708
Qy 149 AlaThrAspProThrProSerValValHisValCysGluSerAspGln-----166
Db 2709 GACAGGACAAAGACCCCGCGCGG-----GGGACCCAGGACCGAGCA 2750
Qy 167 -----AsnGluGluGluGluGluMetGluArgMetLysArgProLysPro 182
Db 2751 GAGGCCCGAAGAGCGGAGCGCGGAGCGCGGTGCCCGGAGCGCGCGCGCGCG 2807

RESULT 11
US-08-949-386-8
; Sequence 8, Application US/08949386
; Patent No. 6090623
; GENERAL INFORMATION:
; APPLICANT: Harpold, Michael
; APPLICANT: Ellis, Steven
; APPLICANT: Williams, Mark

```

```

; APPLICANT: McCue, Ann
; APPLICANT: Gillespie, Alison
; TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: California
; COUNTRY: US
; ZIP: 92101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/949,386
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/290,012
; FILING DATE: 11-AUG-1994
; APPLICATION NUMBER: 08/149,097
; FILING DATE: 5-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/105,536
; FILING DATE: 11-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L.
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 519808
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 238-0999
; TELEFAX: (619) 238-0062
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7175 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 144..6857
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: 1..143
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: 6855..7175
US-08-949-386-8

Alignment Scores:
Pred. No.: 14.1 Length: 7175
Score: 81.00 Matches: 47
Percent Similarity: 33.79% Conservative: 27
Best Local Similarity: 21.46% Mismatches: 75
Query Match: 7.77% Indels: 70
DB: 2 Gaps: 9

US-09-782-953-7 (1-198) x US-08-949-386-8 (1-7175)
Qy 7 SerTyrAsnPheSerLeuLeuAlaCysValAlaAsnAspValPheSerGluSer 26
Db 2232 AACTACACTGCTGGAATGTCTTTCTGCCATCGCTGGACAACCTGGCCACGCCCAA 2291
Qy 27 GluThrArgAlaLysPheGluSerLeuPheArgThrTyrAspLysAspThrPheGln 46
Db 2292 GAGCTGACCAAGGATGAAGAGGAGATGAAGAGCAGCCATCAGAACCTTGCTCTGCAA 2351
Qy 47 TyrPheLysSerPheLysArgValArgLeuAlaAsnPheSerAsnProLeuSerAlaAlaAsp 66
Db 2352 AAGCCAAAGAGTGGCTGAAGTC-----AGCCCCATGTCTGCCCGGAC 2396
Qy 66 -----66
Db 2397 ATCTCCATCGCCGACGACGACAGACACTCGGCCAAGCGCGCTCGGTGGAGACGCG 2456
Qy 67 ---AlaArgLeuArgLeuHis-----LysThrGluPheLeuGlyLysGluMet 81
Db 2457 GCCAGCCAGCTACGCTGCAGAACCTCGGGCCAGCTGCGAGGCGCTGTACAGCGAGATG 2516
Qy 82 -----LysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAla 97
Db 2517 GACCCCGAGGAGCGCTGCGCTTCCGCACTACGCGCCACCTG-----2558
Qy 98 ProProAsnProAspLysGlnPheLeuLeuSerProPro-----110
Db 2559 -----CGGCCCGACATGAAGACGACCTCGACCGCGCGCTGCTGGAGTGGCGCGC 2612
Qy 111 -----AlaSerProProValGlyTyrLysGlnValGluAspAlaThrProValIleAsn 128
Db 2613 GACGGCGCGCGCGCGCGCGTGGAGGCAAGCCGACCTGAGCTCGGAG-----2663
Qy 129 TyrAspLeuLeuTyrAlaIleSerLysLeuGlyProGlyGluLysTyrGluLeuHisAla 148
Db 2664 -----GCCCGCGAGGCGCTCGACCTCGCGCGAGCGCACCCACCGCACCGC 2708
Qy 149 AlaThrAspProThrProSerValValHisValCysGluSerAspGln-----166
Db 2709 GACAGGACAAAGACCCCGCGCGG-----GGGACCCAGGACCGAGCA 2750
Qy 167 -----AsnGluGluGluGluGluMetGluArgMetLysArgProLysPro 182
Db 2751 GAGGCCCGAAGAGCGGAGCGCGGAGCGCGGTGCCCGGAGCGCGCGCGCGCG 2807

```

Db 2352 AAGCCAAAGAGTGGCTGAAGTC-----ACCCCATCTCTGCGCGCAAC 2396
QY 66 ----- 66
Db 2397 ATCTTCATCGCGCGCAGAGAGAACTCGGCAAGCGCGCTCGTGTGGAGCAGCGG 2456
QY 67 --AAlaTgLeuAgtLeuHis-----LysThrGluPheLeuGlyLysGluMet 81
Db 2457 GCCAGCCAGCTACGCGCTCAGAACTCGGCGCCAGCTCGAGCGCGCTGTACAGCGAGATG 2516
QY 82 -----LysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAla 97
Db 2517 GACCCCGAGAGCGGCTCGCTTCCCACTACGCGCCACTG----- 2558
QY 98 ProProAsnProAspLysGlnPheLeuIleSerProPro----- 110
Db 2559 -----CGCCCGCATGAAAGACGACCTGACCGCGCGCTGTGTGTGAGCTGGCGCGC 2612
QY 111 -----AlaSerProProValGlyTrrLysGlnValGluAspAlaThrProValIleAsn 128
Db 2613 GACGCGCGCGCGCGCGCTGTGAGGCAAGCCCGACCTGAGGCTGCGGAG----- 2663
QY 129 TyrAspLeuLeuTyrAlaIleSerLysLeuGlyProGlyGluLysTyrGluLeuHisAla 148
Db 2664 -----GCCCCGAGGCGCTGACCTCGCGCGAGCAGACCGCGCAGCGC 2708
QY 149 AlaThrAspProThrProSerValValAlaHisValCysGluSerAspGlnIle----- 166
Db 2709 GACAAGACACAGACCCCGCGCGG-----GGGACACGAGACGAGCA 2750
QY 167 -----AsnGluGluGluGluGluMetGluArgMetLysArgProLysAspPro 182
Db 2751 GAGGCCCCGAGGCGAGAGCGGCGGAGCCCGGTGCGCGGAGAGCGCGCGCGCG 2807

RESULT 12
US-08-450-562-8
Sequence 8, Application US/08450562
Patent No. 6096514
GENERAL INFORMATION:
APPLICANT: Harpold, Michael
APPLICANT: Ellis, Steven
APPLICANT: Williams, Mark
APPLICANT: McCue, Ann
APPLICANT: Gillespie, Alison
APPLICANT: Feldman, Daniel
APPLICANT: Brenner, Robert
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSER: Brown, Martin, Haller & McClain
STREET: 1660 Union Street
CITY: San Diego
STATE: California
COUNTRY: US
ZIP: 92101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450,562
FILING DATE: 435
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/404,950
FILING DATE: 13-MAR-1995
APPLICATION NUMBER: 08/336,257
FILING DATE: 7-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/314,083

FILING DATE: 28-SEPT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/311,363
FILING DATE: 23-SEPT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/290,012
FILING DATE: 11-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/223,305
FILING DATE: 4-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/193,078
FILING DATE: 07-FEB-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/149,097
FILING DATE: 5-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/105,536
FILING DATE: 11-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/914,231
FILING DATE: 13-JULY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/868,354
FILING DATE: 10-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/06903
FILING DATE: 14-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/745,206
FILING DATE: 15-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/620,250
FILING DATE: 30-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/603,751
FILING DATE: 08-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/482,384
FILING DATE: 02-FEB-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US89/01408
FILING DATE: 04-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/176,899
FILING DATE: 04-APR-1988
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 6362-519812
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 238-0999
TELEFAX: (619) 238-0062
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 7175 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 144..6857
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..143
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 6855..7175
US-08-450-562-8

Alignment Scores:
Pred. No.: 14.1
Length: 7175

Score: 81.00 Matches: 47
Percent Similarity: 33.79% Conservative: 27
Best Local Similarity: 21.46% Mismatches: 75
Query Match: 7.77% Indels: 70
DB: 3 Gaps: 9

US-09-782-953-7 (1-198) x US-08-450-562-8 (1-7175)

```
Oy 7 SerTyrAsnPheserSerLeuLeuAlaCysValAlaAsnAspValPheSerGluSer 26
Db 2232 AACTACACTCTGCTGAATGCTTTCTGGCCATCGCTGTGGACACCTGGCCAAAGCCCAA 2291
Oy 27 GluThrArgAlaLysPheGluSerLeuPheArgThrTyrAspLysAspThrThrPheGln 46
Db 2292 GAGCTGACCAAGGATGAAGAGGAGATGAAGAAGCAGCAATCAGAACTTGCTCTGCAA 2351
Oy 47 TyrPheLysSerPheLysArgValArgileAsnPheserAsnProLeuSerAlaAlaAsp 66
Db 2352 AAGGCCAAGAAGTGGCTGAAGTC-----AGCCCCATGTCTGCGCGGAAC 2396
Oy 66 ----- 66
Db 2397 ATCTCCATCGCGCAGCAGCAGAACTCGGCCAAGCGCGCTCGTGTGGAGCAGCGG 2456
Oy 67 ---AlaArgLeuArgLeuHis-----LysThrGluPheLeuGlyLysGluMet 81
Db 2457 GCCAGCCAGCTACCGCTGTCAGAACCTCGGGGCCAGCTCGCAGGCGCTGTACAGCGAGATG 2516
Oy 82 -----LysLeuTyrPheAlaGlnThrLeuHisLeuGlySerSerHisLeuAla 97
Db 2517 GACCCCGAGGAGCGGCTGGCTTCCCACTACGCGCCACCTG----- 2558
Oy 98 ProAsnProAspLysGlnPheLeuLeuSerProPro----- 110
Db 2559 -----CGGCCGACATGAAGACGACCACTGGACCGCGCTGGTGGAGCTGGCGCG 2612
Oy 111 -----AlaSerProProValGlyTrpLysGlnValGluAspAlaThrProValIleAsn 128
Db 2613 GACGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2663
Oy 129 TyrAspLeuLeuTyrAlaIleSerLysLeuGlyProGlyGluLysTyrGluLeuHisAla 148
Db 2664 -----GCCCGCGAGGCGCTGACCTCGCGCAGGACCAACCGCGCACCGC 2708
Oy 149 AlaThrAspProThrProSerValValHisValCysGluSerAspGlnGlu----- 166
Db 2709 GACAAGCAAGAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2750
Oy 167 -----AsnGluGluGluGluGluMetGluArgMetLysArgProLysPro 182
Db 2751 GAGCGCGCGAGAGCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2807
```

RESULT 13

US-08-984-709A-8
; Sequence 8, Application US/08984709A
; Patent No. 6320032
; GENERAL INFORMATION:
; APPLICANT: Williams, Mark E.
; APPLICANT: Stauderman, Kenneth A.
; APPLICANT: Harpold, Michael M.
; TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Heller Ehrman White & McAuliffe
; STREET: 4250 Executive Square, Suite 700
; CITY: La Jolla
; STATE: California
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/984,709A
; FILING DATE: 02-DEC-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L.
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 24735-9815 (formerly 6362-9815)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 450-8400
; TELEFAX: (619) 587-5360
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7175 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 144..6857
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: 1..143
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: 6855..7175
; US-08-984-709A-8

Alignment Scores:

Pred. No.:	14.1	Length:	7175
Score:	81.00	Matches:	47
Percent Similarity:	33.79%	Conservative:	27
Best Local Similarity:	21.46%	Mismatches:	75
Query Match:	7.77%	Indels:	70
DB:	4	Gaps:	9

US-09-782-953-7 (1-198) x US-08-984-709A-8 (1-7175)

```
Oy 7 SerTyrAsnPheserSerLeuLeuAlaCysValAlaAsnAspValPheSerGluSer 26
Db 2232 AACTACACTCTGCTGAATGCTTTCTGGCCATCGCTGTGGACACCTGGCCAAAGCCCAA 2291
Oy 27 GluThrArgAlaLysPheGluSerLeuPheArgThrTyrAspLysAspThrThrPheGln 46
Db 2292 GAGCTGACCAAGGATGAAGAGGAGATGAAGAAGCAGCAATCAGAACTTGCTCTGCAA 2351
Oy 47 TyrPheLysSerPheLysArgValArgileAsnPheserAsnProLeuSerAlaAlaAsp 66
Db 2352 AAGGCCAAGAAGTGGCTGAAGTC-----AGCCCCATGTCTGCGCGGAAC 2396
Oy 66 ----- 66
Db 2397 ATCTCCATCGCGCAGCAGCAGAACTCGGCCAAGCGCGCTCGTGTGGAGCAGCGG 2456
Oy 67 ---AlaArgLeuArgLeuHis-----LysThrGluPheLeuGlyLysGluMet 81
Db 2457 GCCAGCCAGCTACCGCTGTCAGAACCTCGGGGCCAGCTCGCAGGCGCTGTACAGCGAGATG 2516
Oy 82 -----LysLeuTyrPheAlaGlnThrLeuHisLeuGlySerSerHisLeuAla 97
Db 2517 GACCCCGAGGAGCGGCTGGCTTCCCACTACGCGCCACCTG----- 2558
Oy 98 ProAsnProAspLysGlnPheLeuLeuSerProPro----- 110
Db 2559 -----CGGCCGACATGAAGACGACCACTGGACCGCGCTGGTGGAGCTGGCGCGC 2612
Oy 111 -----AlaSerProProValGlyTrpLysGlnValGluAspAlaThrProValIleAsn 128
Db 2613 GACGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2663
```

```

Oy 129 TyraBpleuLeuTyraHalleseLyLeuGlyProGlyGluuLyTyraHalle 148
Db 2664 -----GCCCGGAGGGCGTGCACCTCCGCGAGGACACCGGACCGC 2708
Oy 149 AlathraBProTrrProSerValValHleValCyeGluSerApGlnu----- 166
Db 2709 GACAGAGACAGACCCCCCGCGC-----GGGACCGAGGACGAGCA 2755
Oy 167 -----AengJugJugJugJugJugJugMetGluArgMetLyArProLyPpo 182
Db 2751 GAGGCCCCGAGGCGGAGAGCGGGAGACCGCGTCCCGGAGGAGCGCGCGCGG 2807

RESULT 14
; Sequence 8, Application US/08450272
; Patent No. 6387696
; GENERAL INFORMATION:
; APPLICANT: Harpold, Michael
; APPLICANT: Ellis, Steven
; APPLICANT: Williams, Mark
; APPLICANT: McCue, Ann
; APPLICANT: Gillespie, Alison
; APPLICANT: Feldman, Daniel
; APPLICANT: Brenner, Robert
; TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: California
; COUNTRY: US
; ZIP: 92101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/450,272
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/404,950
; FILING DATE: 13-MAR-1995
; APPLICATION NUMBER: 08/336,257
; FILING DATE: 7-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/314,083
; FILING DATE: 28-SEPT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/311,363
; FILING DATE: 23-SEPT-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/290,012
; FILING DATE: 11-AUG-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/223,305
; FILING DATE: 4-APR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/193,078
; FILING DATE: 07-FEB-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/149,097
; FILING DATE: 5-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/105,536
; FILING DATE: 11-AUG-1993

```

```

Prior APPLICATION DATA:
APPLICATION NUMBER: 07/914,231
FILING DATE: 13-JULY-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/868,354
FILING DATE: 10-APR-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/06903
FILING DATE: 14-AUG-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/745,206
FILING DATE: 15-AUG-1991
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 6362-519812
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 238-0999
TELEFAX: (619) 238-0062
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 7175 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 144..6857
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..143
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 6855..7175
US-08-450-272-8

Alignment Scores:
Pred. No.: 14.1 Length: 7175
Score: 81.00 Matches: 47
Percent Similarity: 33.79% Conservative: 27
Best Local Similarity: 21.46% Mismatches: 75
Query Match: 7.77% Indels: 70
DB: 4 Gaps: 9

US-09-782-953-7 (1-198) x US-08-450-272-8 (1-7175)
QY 7 SerTyrsenPheSerSerLeuIleAlaCysValAlaAsnAspAspValPheSerGluSer 26
Db 2232 AACTACACTGCTGCAATGCTCTTCTGGCCATCGCTGTGGACAACTGCCAAGCCCAA 2291
QY 27 GluThrArgAlaIysPheGluSerLeuPheArgThrTyraAspLysAspThrThreGln 46
Db 2232 GAGCTGACCAAGATGAAAGAGAGATGGAAGAAAGCAACCAACCAACCTGCTGCA 2351
QY 47 TyrPheLysSerPheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAsp 66
Db 2351 AAGGCCAAGAAAGTGCTGAAGTC-----AGCCCATGTCTGCCCGCAAC 2396
QY 66 ----- 66
Db 2397 ATCTCCATCGCCGCGACAGCAAACTCGGCCAAGCGCGCTCGATGGAGACAGCGG 2456
QY 67 ---AlaArgLeuArgLeuHis-----LysThrGluPheGluGlyLysGluMet 81
Db 2457 GCCACGCAAGCTACGAGCTGCAGAAACCTGCGGACGCTGCGAAGCCCTGTACAGCGAATG 2516
QY 82 -----LysLeuTyrrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAla 97

```

```
Db 2517 GACCCGAGGAGCGCTGGCTTCCCACTAGCGCCACCTG----- 2558
QY 98 ProProAsnProAspLysGlnPheLeuLeuSerProPro----- 110
Db 2559 -----CGCCCGACATGACAGACGACCTGGACCGCGCTGGTGGAGCTGGCGCG 2612
QY 111 -----AlaSerProProValGlyTrpLysGlnValGluAspAlaThrProValIleAsn 128
Db 2613 GAGCGCGCGCGCGCGCTGGAGGCAAGCCCGACCTGAGGCTGGCGAG----- 2663
QY 129 TyrAspLeuLeuTyrAlaIleSerLysLeuGlyProGlyGluLysTyrGluLeuHisAla 148
Db 2664 -----GCCCGCGAGCGCTGCACCTCCGCGCAGGACCCACCGCACCGC 2708
QY 149 AlaThrAspProThrProSerValValHisValCysGluSerAspGlnGlu----- 166
Db 2709 GACAAGGACAAGACCCCGCGCG-----GGGACACGAGACCGAGCA 2750
QY 167 -----AsnGluGluGluGluMetGluArgMetLysArgProLysPro 182
Db 2751 GAGGCGCGAGGCGGAGCGCGCGCTGCTCCCGGAGGAGCGCGCGCGCG 2807
```

RESULT 15

```
US-09-268-163-7
; Sequence 7, Application US/09268163B
; Patent No. 6353091
; GENERAL INFORMATION:
; APPLICANT: Lipecombe, Diane
; TITLE OF INVENTION: HUMAN N-TYPE CALCIUM CHANNEL ISOFORM AND USES THEREOF
; FILE REFERENCE: B1055/7000
; CURRENT APPLICATION NUMBER: US/09/268,163B
; CURRENT FILING DATE: 1999-03-12
; EARLIER APPLICATION NUMBER: US 60/077,901
; EARLIER FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 7177
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: CDS
; LOCATION: 146..6856
US-09-268-163-7
```

Alignment Scores:

Pred. No.:	14.1	Length:	7177
Score:	81.00	Matches:	47
Percent Similarity:	33.79%	Conservative:	27
Best Local Similarity:	21.46%	Mismatches:	75
Query Match:	7.77%	Indels:	70
DB:	4	Gaps:	9

US-09-782-953-7 (1-198) x US-09-268-163-7 (1-7177)

```
QY 7 SerTyrAsnPheSerLeuIleAlaCysValAlaAsnAspValPheSerGluSer 26
Db 2234 AACTACACTGCTGATGCTCTTCTGGCCATCGCTGTGCACAACCTGGCCACGCCCAA 2293
QY 27 GluThrArgAlaLysPheGluSerPheArgThrTyrAspLysAspThrThrPheGln 46
Db 2294 GAGCTGACCAAGGATGAGAGGAGATGGAAGAGCAGCAATCAGAGGCTTCTCTGCA 2353
QY 47 TyrPheLysSerPheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAsp 66
Db 2354 AAGGCCAAGAGAGTGGCTGAAGTC-----AGCCCCATGCTGCCCGCAAC 2398
QY 66 ----- 66
Db 2399 ATCTCCATCGCCCGCAGGAGGAGACTCGGCCAAGCGCGCTCGGTGTGGAGCAGCGG 2458
QY 67 ---AlaArgLeuArgLeuHis-----LysThrGluPheLeuGlyLysGluMet 81
```

```
Db 2459 GCACAGCCAGTACGGCTGCAGAACCTGCGGGCCAGCTGCGAGCGCTGTACAGCGAGATG 2518
QY 82 -----LysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAla 97
Db 2519 GACCCCGAGGAGCGGCTGGCTTGGCACTACGCGCACCTG----- 2560
QY 98 ProProAsnProAspLysGlnPheLeuIleSerProPro----- 110
Db 2561 -----CGCCCGACATGAAGACGACCTGGACCGCGCTGGTGGAGCTGGCGCGC 2614
QY 111 -----AlaSerProProValGlyTrpLysGlnValGluAspAlaThrProValIleAsn 128
Db 2615 GACCGCGCGCGCGCGCTGGAGGCAAGCCCGACCTGAGGCTGGCGAG----- 2665
QY 129 TyrAspLeuLeuTyrAlaIleSerLysLeuGlyProGlyGluLysTyrGluLeuHisAla 148
Db 2666 -----GCCCGCGAGGCGCTGCACCTCCGCGCAGGACCCACCGCACCGC 2710
QY 149 AlaThrAspProThrProSerValValHisValCysGluSerAspGlnGlu----- 166
Db 2711 GACAAGGACAAGACCCCGCGCG-----GGGACACGAGACCGAGCA 2752
QY 167 -----AsnGluGluGluGluMetGluArgMetLysArgProLysPro 182
Db 2753 GAGGCGCGGAGGCGGAGCGCGCGCTGCTCCCGGAGGAGCGCGCGCGCGCG 2809
```

Search completed: December 14, 2002, 22:30:25
Job time : 46.0506 secs

THIS PAGE BLANK (uspro)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2002, 15:19:20 ; Search time 33.2872 Seconds
(without alignments)
13100.560 Million cell updates/sec

Title: US-09-782-953-1
Perfect score: 599
Sequence: 1 gaggtgcaaggagactcca.....gaggaggagaggagagat 599

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 422006 seqs, 364007625 residues

Total number of hits satisfying chosen parameters: 844012

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Pending Patents NA New:*
- 1: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq.*
 - 2: /cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq.*
 - 3: /cgn2_6/ptodata/1/pna/US07_NEW_COMB.seq.*
 - 4: /cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq.*
 - 5: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq.*
 - 6: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq.*
 - 7: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	413.8	69.1	2050	5	US-09-724-676-12756 Sequence 12756, A
2	413.8	69.1	2050	5	US-09-724-676A-12756 Sequence 12756, A
3	413	68.9	1931	5	US-09-724-676-12748 Sequence 12748, A
4	413	68.9	1931	5	US-09-724-676A-12748 Sequence 12748, A
5	413	68.9	2297	5	US-09-724-676-12747 Sequence 12747, A
6	413	68.9	2297	5	US-09-724-676A-12747 Sequence 12747, A
7	404.8	67.6	2348	6	US-10-240-965-60 Sequence 60, Appl
8	349.2	58.3	1893	5	US-09-724-676-12741 Sequence 12741, A
9	349.2	58.3	1893	5	US-09-724-676A-12741 Sequence 12741, A
10	343.8	57.4	1839	5	US-09-724-676-12743 Sequence 12743, A
11	343.8	57.4	1839	5	US-09-724-676A-12743 Sequence 12743, A
12	343.8	57.4	1839	5	US-09-724-676-12745 Sequence 12745, A
13	343.8	57.4	1876	5	US-09-724-676A-12745 Sequence 12745, A
14	340.4	56.8	911	5	US-09-724-676-12749 Sequence 12749, A
15	340.4	56.8	911	5	US-09-724-676A-12749 Sequence 12749, A
16	338.8	56.6	1213	5	US-09-724-676-12740 Sequence 12740, A
17	338.8	56.6	1213	5	US-09-724-676A-12740 Sequence 12740, A
18	338	56.4	1094	5	US-09-724-676-12750 Sequence 12750, A
19	338	56.4	1094	5	US-09-724-676A-12750 Sequence 12750, A
20	274.2	45.8	1056	5	US-09-724-676-12752 Sequence 12752, A
21	274.2	45.8	1056	5	US-09-724-676A-12752 Sequence 12752, A
22	268.8	44.9	1002	5	US-09-724-676-12744 Sequence 12744, A
23	268.8	44.9	1002	5	US-09-724-676A-12744 Sequence 12744, A
24	268.8	44.9	1039	5	US-09-724-676-12746 Sequence 12746, A
25	268.8	44.9	1039	5	US-09-724-676A-12746 Sequence 12746, A
26	264.4	44.1	442	5	US-09-513-999C-1772 Sequence 1772, Ap

27	203.8	34.0	2411	5	US-09-724-676-15329	Sequence 15329, A
28	203.8	34.0	2411	5	US-09-724-676A-15329	Sequence 15329, A
29	197.6	33.0	3159	6	US-10-290-438-1	Sequence 1, Appli
30	181.8	30.4	615	6	US-10-290-438-8	Sequence 8, Appli
C 31	153	25.5	1021	6	US-10-290-438-3	Sequence 3, Appli
C 32	144.6	24.1	446	6	US-10-203-138A-2101	Sequence 2101, Ap
33	143.6	24.0	1577	5	US-09-724-676-12755	Sequence 12755, A
34	143.6	24.0	1577	5	US-09-724-676A-12755	Sequence 12755, A
35	143.6	24.0	1943	5	US-09-724-676-12752	Sequence 12752, A
36	143.6	24.0	1943	5	US-09-724-676A-12752	Sequence 12752, A
C 37	77.4	12.9	123	6	US-10-203-138A-5980	Sequence 5980, Ap
C 38	75.6	12.6	486	6	US-10-203-138A-827	Sequence 827, App
39	72.2	12.1	305	5	US-09-724-676-12751	Sequence 12751, A
40	72.2	12.1	305	5	US-09-724-676A-12751	Sequence 12751, A
41	71	11.9	557	5	US-09-724-676-12753	Sequence 12753, A
42	71	11.9	557	5	US-09-724-676A-12753	Sequence 12753, A
C 43	70.6	11.8	85	6	US-10-203-138A-7216	Sequence 7216, Ap
44	68.6	11.5	740	5	US-09-724-676-12754	Sequence 12754, A
45	68.6	11.5	740	5	US-09-724-676A-12754	Sequence 12754, A

ALIGNMENTS

RESULT 1
US-09-724-676-12756
; Sequence 12756, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 12756
; LENGTH: 2050
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (7)..(7)
; OTHER INFORMATION: n is a,c,g, or t
; NAME/KEY: misc feature
; LOCATION: (47)..(47)
; OTHER INFORMATION: n is a,c,g, or t
US-09-724-676-12756

Query Match 69.1%; Score 413.8; DB 5; Length 2050;
Best Local Similarity 85.1%; Pred. No. 1e-119;
Matches 474; Conservative 0; Mismatches 82; Indels 1; Gaps 1;

Qy	39	GAGCGAGTGGTTCGTTAAGCGTCGCCCGTGAAGAGCAGATGATTTTAGGACTTTA	98
Db	141	GACTGCGGGTCTGTGACGCTTTCACTGAAGAAGCAAGATGCAATTTAGAACTTTA	200
Qy	99	GCTACAAATTTAGCTCCCTGATGCTGTGCGCAACGATGATGCTTCAGGAAAGTG	158
Db	201	ACTACAGTTTGTAGCTCCCTGATGCTGTGCGCAACGATGATGCTTCAGGAAAGTG	260
Qy	159	AGACGAGGGCCAAATTTGAATCCCTCTTCAGAACATATGACAAAGGACACCACCTTCAGT	218
Db	261	AAACGAGGGCCAAATTTGAGTCCCTCTTTAGGAGGTATGACAGGACATCACCTTCAGT	320
Qy	219	ATTTTAAGAGCTTCAACAGTGTCCGGATAAATTCAGCAACCCCTTATCTGCAGCCGATG	278
Db	321	ATTTTAAGAGCTTCAACAGGATGAGTAATTAATTCAGCAACCCCTTCTCCGACGAGATG	380
Qy	279	CCAGGCTGCGGCTGCACAGACCCAGTTCCTGGGGAGGAAATGAAAGTTGTATTTGCTC	338
Db	381	CCAGGCTCCAGCTGCATAGAGTGTCTCTGGGAAAGAAATGAAGTTATATTTTGCTC	440

Db 326 CTTACACATAGGAAGCTCACACCTGGCTCCGCCAAATCCAGACAAGCAGTTTCTGATCTC 385
QY 402 CCTCCGGGCTCTCTCCCGTTGGTGGAAACAAGTAGAAGATGCCACCCCGTCATAAA 461
Db 386 CCTCCCGGCTCTCCCGCAGTGGATGAAACAAGTAGAAGATGCCACCCCGTCATAAA 445
QY 462 TTACGATCTTTATATGCGATCTCCAAAGCTGGGGCCAGGAGAAAGTATGAATTCGATGC 521
Db 446 CTATGATCTCTTATATGCGATCTCCAAAGCTGGGGCCAGGAGAAAGTATGAATTCGATGC 505
QY 522 AGCGACAGACACCACTCCCAAGTGTGGTCCACGTGTGAGAGTGCACCAAGAGATGA 581
Db 506 AGCGACTGACACCACTCCCAAGTGTGGTCCATGTATGTGAGAGTGCATCAAGAGAAGGA 565
QY 582 GGAGGAAGAGGAA 594
Db 566 GGAAGAAGAGGAA 578

RESULT 4

US-09-724-676A-12748
; Sequence 12748, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724, 676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12748
; LENGTH: 1931
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (11)..(11)
; OTHER INFORMATION: n is a,c,g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (18)..(18)
; OTHER INFORMATION: n is a,c,g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (23)..(23)
; OTHER INFORMATION: n is a,c,g, or t
US-09-724-676A-12748

Query Match 68.9%; Score 413; DB 5; Length 1931;
Best Local Similarity 85.4%; Pred. No. 1.8e-119;
Matches 472; Conservative 0; Mismatches 80; Indels 1; Gaps 1;

QY 43 GAGTCGTTCTGTTAAGCGTCTGCCCGCTGAAAGAGCAAGATGATTTTAGGACTTTAGCTA 102
Db 26 GCGTGGTCTGTAGCGCTTTCACTGTAAGAAAGCAAGATGATTTTAGAACTTTAACTA 85
QY 103 CAATTTTAGCTCCCTGATTCCTGTCGCAACAGATGATGCTTCAGCGAAGTGAGAC 162
Db 86 CAGTTTGTAGTCCCTGATTCCTGTCGCAACAGATGATGATCTTCAGCGAAGTGAAAC 145
QY 163 CAGGGCCAAATTGAATTCCTCTTCAGAACATATGACAGGACACACCTTCAGTATTT 222
Db 146 CAGGGCCAAATTGAGTTCCTCTTTAGGACGTATGACAGGACATCACCTTCAGTATTT 205
QY 223 TAAGAGCTTCAACAGCTCCCGATTAACCTTCAGCAACCCCTTATCTGACGCGATGCCAG 282
Db 206 TAAGAGCTTCAACAGCTCCCGATTAACCTTCAGCAACCCCTTCTCCGACGAGATGCCAG 265
QY 283 GCTGCGGCTGCAACAGACCCAGTTCTCGGGGAAGGAAATGAAGTTGTATTTTGTCTCAGAC 342
Db 266 GCTCCAGCTGCATAGACTGAGTTCTGGGAAGGAAATGAAGTTGTATTTTGTCTCAGAC 325
QY 343 TTACACATAGGAAGTTCAACACCTGGCTCCGCC-AAATCCCGACAAACAGTTCTCTCATCTC 401

Db 326 CTTACACATAGGAAGCTCACACCTGGCTCCGCCAAATCCAGACAAGCAGTTTCTGATCTC 385
QY 402 CCTCCGGGCTCTCTCCCGTTGGTGGAAACAAGTAGAAGATGCCACCCCGTCATAAA 461
Db 386 CCTCCCGGCTCTCCCGCAGTGGATGAAACAAGTAGAAGATGCCACCCCGTCATAAA 445
QY 462 TTACGATCTTTATATGCGATCTCCAAAGCTGGGGCCAGGAGAAAGTATGAATTCGATGC 521
Db 446 CTATGATCTCTTATATGCGATCTCCAAAGCTGGGGCCAGGAGAAAGTATGAATTCGATGC 505
QY 522 AGCGACAGACACCACTCCCAAGTGTGGTCCACGTGTGAGAGTGCACCAAGAGATGA 581
Db 506 AGCGACTGACACCACTCCCAAGTGTGGTCCATGTATGTGAGAGTGCATCAAGAGAAGGA 565
QY 582 GGAGGAAGAGGAA 594
Db 566 GGAAGAAGAGGAA 578

RESULT 5

US-09-724-676-12747
; Sequence 12747, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724, 676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12747
; LENGTH: 2297
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (11)..(11)
; OTHER INFORMATION: n is a,c,g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (18)..(18)
; OTHER INFORMATION: n is a,c,g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (23)..(23)
; OTHER INFORMATION: n is a,c,g, or t
US-09-724-676-12747

Query Match 68.9%; Score 413; DB 5; Length 2297;
Best Local Similarity 85.4%; Pred. No. 1.9e-119;
Matches 472; Conservative 0; Mismatches 80; Indels 1; Gaps 1;

QY 43 GAGTCGTTCTGTTAAGCGTCTGCCCGCTGAAAGAGCAAGATGATTTTAGGACTTTAGCTA 102
Db 26 GCGTGGTCTGTAGCGCTTTCACTGTAAGAAAGCAAGATGATTTTAGAACTTTAACTA 85
QY 103 CAATTTTAGCTCCCTGATTCCTGTCGCAACAGATGATGCTTCAGCGAAGTGAGAC 162
Db 86 CAGTTTGTAGTCCCTGATTCCTGTCGCAACAGATGATGATCTTCAGCGAAGTGAAAC 145
QY 163 CAGGGCCAAATTGAATTCCTCTTCAGAACATATGACAGGACACACCTTCAGTATTT 222
Db 146 CAGGGCCAAATTGAGTTCCTCTTTAGGACGTATGACAGGACATCACCTTCAGTATTT 205
QY 223 TAAGAGCTTCAACAGCTCCCGATTAACCTTCAGCAACCCCTTATCTGACGCGATGCCAG 282
Db 206 TAAGAGCTTCAACAGCTCCCGATTAACCTTCAGCAACCCCTTCTCCGACGAGATGCCAG 265
QY 283 GCTGCGGCTGCAACAGACCCAGTTCTCGGGGAAGGAAATGAAGTTGTATTTTGTCTCAGAC 342
Db 266 GCTCCAGCTGCATAGACTGAGTTCTGGGAAGGAAATGAAGTTGTATTTTGTCTCAGAC 325

QY 343 TTACACATAGAGAAATTGACACCTGCTCCGCC-AATCCCGACAAACAGTTCTCATCTC 401
DB 326 CTACACATAGAGAAAGCTCACACCTGCTCCGCCAAATCCAGACAAAGCAGTTCTCATCTC 385
QY 402 CCTCCCGGCTCTCTCCGCTGCTGAGAAACAAAGTAGAAGTCCACCCCTCATATAA 461
DB 386 CCTCCCGGCTCTCTCCGCTGCTGAGAAACAAAGTAGAAGTCCACCCCTCATATAA 445
QY 462 TTACGATCTTTATATGCTCATCTCCAAAGTGGGCCAGAGAGAGATGATGATGATG 521
DB 446 CTATGATCTCTTATATGCTCATCTCCAAAGTGGGCCAGAGAGAGATGATGATGATG 505
QY 522 AGGACAGACACCACTCCGAGTGTGTGTGTCCAGCTGTGTGAGAGTGAACCAAGATGA 581
DB 506 AGGACTGACACCACTCCGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 565
QY 582 GGAGAGAGAGAA 594
DB 566 GGAGAGAGAGAA 578

RESULT 6
US-09-724-676A-12747
Sequence 12747, Application US/09724676A
GENERAL INFORMATION:
APPLICANT: Compugen LTD
TITLE OF INVENTION: Variants of alternative splicing
FILE REFERENCE: 129181.4 Compugen
CURRENT APPLICATION NUMBER: US/09/724,676A
CURRENT FILING DATE: 2000-11-28
NUMBER OF SEQ ID NOS: 97222
SOFTWARE: PatentIn version 3.2
SEQ ID NO 12747
LENGTH: 2297
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (11)..(11)
OTHER INFORMATION: n is a,c,g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (18)..(18)
OTHER INFORMATION: n is a,c,g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (23)..(23)
OTHER INFORMATION: n is a,c,g, or t
US-09-724-676A-12747

Query Match 68.9%; Score 413; DB 5; Length 2297;
Best Local Similarity 85.4%; Pred. No. 1.9e-119;
Matches 472; Conservative 0; Mismatches 80; Indels 1; Gaps 1;

QY 43 GAGTCGTTGTTAAGCGTCTGCCCCGCGGAAAGCAAGATATTTTGGGAGCTTTACTA 102
DB 26 GCGTGGGCTGTAGCGCTTTCACCTGTAGAAAGCAAGATGATTTTAACTTA 85
QY 103 CAATTTAGCTCCCTGATGCTGTGTGTGCAAAAGATGATCTTTACGGAAGTGAAC 162
DB 86 CAGTTTAGCTCCCTGATGCTGTGTGTGCAAAAGATGATCTTTACGGAAGTGAAC 145
QY 163 CAGGCGCAATTTGAATCCCTCTTCAGAACATATGACAAGACACACCTTCCAGTATTT 222
DB 146 CAGGCGCAATTTGAATCCCTCTTCAGAACATATGACAAGACATCACTTTCAGTATTT 205
QY 223 TAAAGAGCTTCAAAAGTGTCCGATTAATCTTCAGCAACCCCTTATGAGCGGATGCGAG 282
DB 206 TAAAGAGCTTCAAAAGTGTCCGATTAATCTTCAGCAACCCCTTCTCGAGAGATGCGAG 265
QY 263 GCTGGGGCTGCAAGACCGAGTCTCTGGGGAAGAAATGAAGTTGATTTTGTCTAGAC 342
DB 266 GCTCGAGCTGCAATGAGTCTCTGGGGAAGAAATGAAGTTGATTTTGTCTAGAC 325

QY 343 TTACACATAGAGAAATTGACACCTGCTCCGCC-AATCCCGACAAACAGTTCTCATCTC 401
DB 326 CTACACATAGAGAAAGCTCACACCTGCTCCGCCAAATCCAGACAAAGCAGTTCTCATCTC 385
QY 402 CCTCCCGGCTCTCTCCGCTGCTGAGAAACAAAGTAGAAGTCCACCCCTCATATAA 461
DB 386 CCTCCCGGCTCTCTCCGCTGCTGAGAAACAAAGTAGAAGTCCACCCCTCATATAA 445
QY 462 TTACGATCTTTATATGCTCATCTCCAAAGTGGGCCAGAGAGAGATGATGATGATG 521
DB 446 CTATGATCTCTTATATGCTCATCTCCAAAGTGGGCCAGAGAGAGATGATGATGATG 505
QY 522 AGGACAGACACCACTCCGAGTGTGTGTGTCCAGCTGTGTGAGAGTGAACCAAGATGA 581
DB 506 AGGACTGACACCACTCCGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 565
QY 582 GGAGAGAGAGAA 594
DB 566 GGAGAGAGAGAA 578

RESULT 7
US-10-240-965-60
Sequence 60, Application US/10240965
GENERAL INFORMATION:
APPLICANT: INCYTE GENOMICS, INC.
APPLICANT: SHIFMAN, Dov
APPLICANT: SOMOGYI, Roland
APPLICANT: LAWN, Richard M.
APPLICANT: SEITHAMER, Jeffrey J.
APPLICANT: PORTER, Gordon J.
APPLICANT: MIKITA, Thomas
TITLE OF INVENTION: GENES EXPRESSED IN FOAM CELL DIFFERENTIATION
FILE REFERENCE: PA-0025 PCT
CURRENT APPLICATION NUMBER: US/10/240,965
CURRENT FILING DATE: 2002-10-04
PRIOR FILING DATE: 2000-04-05
NUMBER OF SEQ ID NOS: 276
SOFTWARE: PERL Program
SEQ ID NO 60
LENGTH: 2348
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No: 042176.5
US-10-240-965-60

Query Match 67.6%; Score 404.8; DB 6; Length 2348;
Best Local Similarity 84.1%; Pred. No. 7.3e-117;
Matches 469; Conservative 0; Mismatches 87; Indels 2; Gaps 1;

QY 39 GAGCGAGCTGTTGTTAAGCGTCTGCCCCGCGGAAAGCAAGATATTTTGGGAGCTTTA 98
DB 110 GACTGGGTGGTGTGAGCGCTTTCACCTGTAGAAAGCAAGATGATTTTAACTTA 169
QY 99 GCTACATTTTAACTCCCTGATGCTGTGTGTGCAAAAGATGATGCTTTACGGAAGT 158
DB 170 ACTACATTTTAACTCCCTGATGCTGTGTGTGCAAAAGATGATGCTTTACGGAAGT 229
QY 159 AGACAGGCGCAATTTGAATCCCTCTTCAGAACATATGACAAGACACACCTTCCAGT 218
DB 230 AAACAGGCGCAATTTGAATCCCTCTTCAGAACATATGACAAGACATCACTTTCACT 289
QY 219 ATTTAAGAGCTTCAAAAGTGTCCGATTAATCTTCAGCAACCCCTTATGAGCGGATG 278
DB 290 ATTTAAGAGCTTCAAAAGTGTCCGATTAATCTTCAGCAACCCCTTCTCGAGAGATG 349
QY 279 CAGGCTGCGCTGCAAGACCGAGTCTCTGGGGAAGAAATGAAGTTGATTTTGTCTC 338

Db 350 *CCAGGCTCCAGCTGCATAGAGCTAGTTTCTGGGAAAGAAATGAAGTTATATATTTTGCTC 409
QY 339 AGACTTTACATAGGAAGTTTCACACCTGG--CTCGGCAATCCCGACAAACAGTTCTTC 396
Db 410 AGACCTTACATAGGAAGTTTCACACCTGGGCTCGGCAATCCAGAACAGAGTTTCTG 469
QY 397 ATCTCCCTCCGCTCTCTCTCCGCTGGCTGGGAAACAAAGTAGAGATGCCACCCCGCTC 456
Db 470 ATCTCCCTCCGCTCTCTCCGCTGGGATGGAAACAAAGTGAAGATGCCACCCAGTC 529
QY 457 ATAAATATAGATCTTTTATATGCAATCTCCAGCTGGGCTGGGCAAGAGATGTAAGTCTG 516
Db 530 ATAAATATAGATCTTTTATATGCAATCTCCAGCTGGGCTGGGCAAGAGATGTAAGTCTG 589
QY 517 CATGAGGAGAGACACCACTCCAGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 576
Db 590 CACGAGGAGAGACACCACTCCAGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 649
QY 577 AATGAGGAGAGAGAGAA 594
Db 650 AAGGAGAGAGAGAGAA 667

RESULT 8

US-09-724-676-12741

; Sequence 12741, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12741
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676-12741

Query Match 58.3%; Score 349.2; DB 5; Length 1893;
Best Local Similarity 80.9%; Pred. No. 1.9e-99;
Matches 419; Conservative 0; Mismatches 98; Indels 1; Gaps 1;

QY 78 AGAATGATTTTAGGACCTTTAGCTACAAATTTAGCTCCCTGATGTTGTTGGTGGCAACG 137
Db 23 AGATGAGGAGGTGGACCTCGAGACCTGCCAGCGCCAAATTTAGTCCCTTTTAGACGTATG 82
QY 138 ATGATGCTTTTCAGCGAAGTGAAGACCAAGGCGGCAAAATTTGAATCCCTCTTCAGAACATATG 197
Db 83 CGCGCGTGTTCGTGGAGCGGCTGTGCGGCGCAAAATTTAGTCCCTTTTAGACGTATG 142
QY 198 ACAAGGACACCACTTCCAGTATTTTAGAGCTTCAAAAGCTTCGCGATAAACTTTCAGCA 257
Db 143 ACAAGGACATCACTTTTCAGTATTTTAGAGCTTCAAAAGCTTCAGAACATATG 202
QY 258 ACCCTTTATCTGAGCGGATGCCAGGCTGCCAGCTGCCAGCGAGTTCCTCGGGAAGG 317
Db 203 ACCCTTTCTCGGAGAGTGCAGCTGCCAGGCTGCCAGCTGCCAGCTGCCAGCTGCCAGG 262
QY 318 AAATCAAGTTGATTTGCTCAGACTTTACATAGGAAGTTTCACTGGCTCCGCC-A 376
Db 263 AAATGAAGTTATATTTTGTCTCAGACCTTACATAGGAAGTTCACACTGGCTCCGCCAA 322
QY 377 ATCCGCAAAACAGTTCTCATCTCCCTCCGCGCTCTCTCCGTTGGCTGGGAACAAG 436
Db 323 ATCCAGACAGCAGTTTCTGATCTCCCTCCGCTCTCCGCGAGTGGGATGGAAACAAG 382
QY 437 TAGAAGATGCCACCCCGCTCATAAATTTACGATCTTTTATATGCCATCTCCAAGCTGGGGC 496
Db 383 TGAAGATGGCGACCCAGCTCATAAATTTATGATCTTTATATGCCATCTCCAAGCTGGGGC 442
QY 497 CAGGAGAGAGTATGAATCTGCATGCGGAGCAGACACCACTCCAGTGTGGTGGTCCAGC 556
Db 443 CAGGGGAAAGTATGAATTTGCACGCGACTGACACCACTCCAGCTCCAGCTGGTGGTCCATG 502
QY 557 TGTGTGAGGTGACCAAGAGAGATGAGGAGAGAGAGAA 594
Db 503 TATGTGAGGTGATCAAGAGAGAGAGAGAGAGAA 540

RESULT 10

US-09-724-676-12743

; Sequence 12743, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen

CURRENT APPLICATION NUMBER: US/09/724.676
 ; CURRENT FILING DATE: 2000-11-28
 ; NUMBER OF SEQ ID NOS: 97222
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 12743
 ; LENGTH: 1839
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-724-676-12743

Query Match 57.4%; Score 343.8; DB 5; Length 1839;
 Best Local Similarity 88.9%; Pred. No. 9.1e-98;
 Matches 383; Conservative 0; Mismatches 47; Indels 1; Gaps 1;

QY 165 GGGCCAAATTGATCCCTCTTCAGAACATATGACAGAGACACCCCTTCAGATATTTTA 224
 Db 56 GGGCCAAATTGATCCCTCTTCAGAACATATGACAGAGACACCCCTTCAGATATTTTA 115
 QY 225 AAGAGCTTCAAGGTGTCGGGATTAACCTTACAGCAACCCCTTATCTGACGCCGATGCCAGGC 284
 Db 116 AAGAGCTTCAAGGTGTCGGGATTAACCTTACAGCAACCCCTTATCTGACGCCGATGCCAGGC 175
 QY 285 TCGGCTGCACAAAGACCGAGTTCTTGGGGAAGAAATGAATTGTTTGTCTCAGACTT 344
 Db 176 TCGAGCTGCATAGACTGAGTTCTTGGGGAAGAAATGAATTGTTTGTCTCAGACTT 235
 QY 345 TACACATAGGAAGTTTACACACTTGGCTCCGCC- AATCCGCAAAAGATTCTCATCTCCC 403
 Db 236 TACACATAGGAAGTTTACACACTTGGCTCCGCCAATCCAGACAGATTCTCATCTCCC 295
 QY 404 CTCGGGCTCTCTCCGCTTGGCTGGAACAAAGTAGAAGATGCCACCCCGTCATTAATT 463
 Db 296 CTCGGGCTCTCTCCGCTTGGCTGGAACAAAGTAGAAGATGCCACCCCGTCATTAATT 355
 QY 464 ACGATCTTTATATGCGATCTTCCAGCTGGGCCAGAGAGAAATGAATGATGCAATGCGAG 523
 Db 356 ATGATCTTTATATGCGATCTTCCAGCTGGGCCAGAGAGAAATGAATGATGCAATGCGAG 415
 QY 524 CGACAGACCACTCCCAAGTGTGTGTCACAGTGTGAGAGTGAACCAAGAAATGAG 583
 Db 416 CGACTGACACCACTCCCAAGTGTGTGTCATGTATGTGAGATGATCAAGAGAAAGAGG 475
 QY 584 AGAAGAGGAA 594
 Db 476 AAGAAGAGGAA 486

RESULT 11
 US-09-724-676A-12743
 ; Sequence 12743, Application US/09724676A
 ; GENERAL INFORMATION:
 ; APPLICANT: Compugen LTD
 ; TITLE OF INVENTION: Variance of alternative splicing
 ; FILE REFERENCE: 129181.4 Compugen
 ; CURRENT APPLICATION NUMBER: US/09/724.676A
 ; CURRENT FILING DATE: 2000-11-28
 ; NUMBER OF SEQ ID NOS: 97222
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 12743
 ; LENGTH: 1839
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-724-676A-12743

Query Match 57.4%; Score 343.8; DB 5; Length 1839;
 Best Local Similarity 88.9%; Pred. No. 9.1e-98;
 Matches 383; Conservative 0; Mismatches 47; Indels 1; Gaps 1;

QY 165 GGGCCAAATTGATCCCTCTTCAGAACATATGACAGAGACCACTTCAGATATTTTA 224
 Db 56 GGGCCAAATTGATCCCTCTTCAGAACATATGACAGAGACCACTTCAGATATTTTA 115
 QY 225 AAGAGCTTCAAGGTGTCGGGATTAACCTTACAGCAACCCCTTATCTGACGCCGATGCCAGGC 284

Db 116 AAGAGCTTCAAGGTGTCGGGATTAACCTTACAGCAACCCCTTCAGATATTTTA 175
 QY 285 TCGGCTGCACAAAGACCGAGTTCTTGGGGAAGAAATGAATTGTTTGTCTCAGACTT 344
 Db 176 TCGAGCTGCATAGACTGAGTTCTTGGGGAAGAAATGAATTGTTTGTCTCAGACTT 235
 QY 345 TACACATAGGAAGTTTACACACTTGGCTCCGCC- AATCCGCAAAAGATTCTCATCTCCC 403
 Db 236 TACACATAGGAAGTTTACACACTTGGCTCCGCCAATCCAGACAGATTCTCATCTCCC 295
 QY 404 CTCGGGCTCTCTCCGCTTGGCTGGAACAAAGTAGAAGATGCCACCCCGTCATTAATT 463
 Db 296 CTCGGGCTCTCTCCGCTTGGCTGGAACAAAGTAGAAGATGCCACCCCGTCATTAATT 355
 QY 464 ACGATCTTTATATGCGATCTTCCAGCTGGGCCAGAGAGAAATGAATGAACTGATGAG 523
 Db 356 ATGATCTTTATATGCGATCTTCCAGCTGGGCCAGAGAGAAATGAATGAACTGATGAG 415
 QY 524 CGACAGACCACTCCCAAGTGTGTGTCACAGTGTGAGAGTGAACCAAGAAATGAG 583
 Db 416 CGACTGACACCACTCCCAAGTGTGTGTCATGTATGTGAGATGATCAAGAGAAAGAGG 475
 QY 584 AGAAGAGGAA 594
 Db 476 AAGAAGAGGAA 486

RESULT 12
 US-09-724-676-12745
 ; Sequence 12745, Application US/09724676
 ; GENERAL INFORMATION:
 ; APPLICANT: Compugen LTD
 ; TITLE OF INVENTION: Variance of alternative splicing
 ; FILE REFERENCE: 129181.4 Compugen
 ; CURRENT APPLICATION NUMBER: US/09/724.676
 ; CURRENT FILING DATE: 2000-11-28
 ; NUMBER OF SEQ ID NOS: 97222
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 12745
 ; LENGTH: 1876
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-724-676-12745

Query Match 57.4%; Score 343.8; DB 5; Length 1876;
 Best Local Similarity 88.9%; Pred. No. 9.2e-98;
 Matches 383; Conservative 0; Mismatches 47; Indels 1; Gaps 1;

QY 165 GGGCCAAATTGATCCCTCTTCAGAACATATGACAGAGACCACTTCAGATATTTTA 224
 Db 93 GGGCCAAATTGATCCCTCTTCAGAACATATGACAGAGACCACTTCAGATATTTTA 152
 QY 225 AAGAGCTTCAAGGTGTCGGGATTAACCTTACAGCAACCCCTTATCTGACGCCGATGCCAGGC 284
 Db 153 AAGAGCTTCAAGGTGTCGGGATTAACCTTACAGCAACCCCTTATCTGACGCCGATGCCAGGC 212
 QY 285 TCGGCTGCACAAAGACCGAGTTCTTGGGGAAGAAATGAATTGTTTGTCTCAGACTT 344
 Db 213 TCGAGCTGCATAGACTGAGTTCTTGGGGAAGAAATGAATTGTTTGTCTCAGACTT 272
 QY 345 TACACATAGGAAGTTTACACACTTGGCTCCGCC- AATCCGCAAAAGATTCTCATCTCCC 403
 Db 273 TACACATAGGAAGTTTACACACTTGGCTCCGCCAATCCAGACAGATTCTCATCTCCC 332
 QY 404 CTCGGGCTCTCTCCGCTTGGCTGGAACAAAGTAGAAGATGCCACCCCGTCATTAATT 463
 Db 333 CTCGGGCTCTCTCCGCTTGGCTGGAACAAAGTAGAAGATGCCACCCCGTCATTAATT 392
 QY 464 ACGATCTTTATATGCGATCTTCCAGCTGGGCCAGAGAGAAATGAATGAACTGATGAG 523
 Db 393 ATGATCTTTATATGCGATCTTCCAGCTGGGCCAGAGAGAAATGAATGAACTGATGAG 452

QY 524 CGACAGACACCACTCCCAAGTGTGGTCCAGTGTGTGAGAGTGAACCAAGAGAAATGAGG 583
Db 453 CGACTGACACCACTCCCAAGTGTGGTCCATGTATGTGAGAGTGAATCAAGAGAGGAGG 512
QY 584 AGGAAGAGGAA 594
Db 513 AAGAAGAGGAA 523

RESULT 13

US-09-724-676A-12745
; Sequence 12745, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12745
; LENGTH: 1876
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676A-12745

Query Match 57.4%; Score 343.8; DB 5; Length 1876;
Best Local Similarity 88.9%; Pred. No. 9.2e-98;
Matches 383; Conservative 0; Mismatches 47; Indels 1; Gaps 1;
QY 165 GGGCCAAATTTGAATCCCTCTTCAGACATATGACAGGACACCACTTCAGTATTTTA 224
Db 93 GGGCCAAATTTGATCCCTCTTTAGGACGTATGACAGGACATCACCTTTTCAGTATTTTA 152
QY 225 AGAGCTTCAAAACGTTCCGGATAACTTCAGCAACCCCTTATCTGACGCGGATGCCAGGC 284
Db 153 AGAGCTTCAAAACGAGTCAGATAAACTTCAGCAACCCCTTCTCGGACGAGATGCCAGGC 212
QY 285 TCGCGCTGCACAGACCGAGTCTCGGGAGGAATGAAGTTGTATTTTGTCTCAGACTT 344
Db 213 TCCAGCTGCATAAGACTGAGTTCTTGGGAAAGGAATGAAGTTATATTTTGTCTCAGACTT 272
QY 345 TACACATAGGAAGTTACACTGTGCTCCGCC-AATCCGACAAACAGTTCTCTCATCTCCC 403
Db 273 TACACATAGGAAGTTACACTGTGCTCCGCCAAATCCAGACAGCAGTTCTCATCTCCC 332
QY 404 CTCGGGCTCTCTCTCCGTTGGCTGGAAACAAGTAGAAGATGCCACCCCGTCAATAAT 463
Db 333 CTCGGGCTCTCTCCGAGTGGGATGAAACAAGTGAAGATGCCACCCAGTCTAACT 392
QY 464 ACGATCTTTATATGCCATCTCCAGTGGGGCCAGGAGAAATGAATGCAATGCGAG 523
Db 393 ATGATCTTTATATGCCATCTCCAGTGGGGCCAGGAGAAATGAATGCGAG 452
QY 524 CGACAGACCACTCCCAAGTGTGGTGTCCACGTGTGAGAGTGAACCAAGAGAAATGAGG 583
Db 453 CGACTGACACCACTCCCAAGTGTGGTGTCCATGTATGTGAGAGTGAACCAAGAGAGGAGG 512
QY 584 AGGAAGAGGAA 594
Db 513 AAGAAGAGGAA 523

RESULT 14

US-09-724-676-12749
; Sequence 12749, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222

; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12749
; LENGTH: 911
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (11)-(11)
; OTHER INFORMATION: n is a,c,g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (18)-(18)
; OTHER INFORMATION: n is a,c,g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (23)-(23)
; OTHER INFORMATION: n is a,c,g, or t
US-09-724-676-12749

Query Match 56.8%; Score 340.4; DB 5; Length 911;
Best Local Similarity 84.5%; Pred. No. 8e-97; Mismatches 0; Indels 1; Gaps 1;
Matches 394; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 43 GAGTCGTTCTTAAGCGTCTGCCCGCTGAAAGAGCAGAAATGATTTTAGGCACTTTAGCTA 102
Db 26 GCGTGGTCTGTAGCGCTTTCACTGTGAAGAGCAAGATGCATTTTAGAACTTTTAACTA 85
QY 103 CAATTTAGCTCCCTGATTTGTTGTGTCGCAAAACGATGATGTTCTTACGCGAAAGTGAGAC 162
Db 86 CAGTTTATAGCTCCCTGATTTGCTGTGTCGCAAAACGATGATGATCTTTCAGCGAAAGTGAAAC 145
QY 163 CAGGGCCAAATTTGAATCCCTCTTCAGACATATGACAGGACACCACTTCAGTATTTT 222
Db 146 CAGGGCCAAATTTGATTCCTCTTTAGGACGTATGACAGGACATCACCTTTTCAGTATTT 205
QY 223 TAAGAGCTTCAAAACGTTCCGGATAACTTCAGCAACCCCTTATCTGACGCGATGCCAG 282
Db 206 TAAGAGCTTCAAAACGAGTCAGATAAACTTCAGCAACCCCTTCTCCGACGAGATGCCAG 265
QY 283 GCTGCGGCTGCACAGACCGAGTCTCTGGGGAAGAAATGAAGTTGTATTTTGTCTCAGAC 342
Db 266 GCTCAGCTGCATAGACTGAGTTCTGGGAAAGGAATGAAGTTATATTTTGTCTCAGAC 325
QY 343 TTACACATAGGAAGTTACACTGTGCTCCGCC-AATCCGACAAACAGTTCTCTCATCTC 401
Db 326 CTTCACATAGGAAGTCTACACTGTGCTCCGCCAAATCCAGACAGCAGTTTCTGATCTC 385
QY 402 CCTCCGGCTCTCTCTCCGTTGGCTGGAAACAAGTAGAAGATGCCACCCCGTCAATAA 461
Db 386 CCTCCCGCTCTCCGCGAGTGGGATGAAACAAGTGAAGATGCCACCCAGTCATAA 445
QY 462 TPAAGATCTTTTATATGCCATCTCCAAAGCTGGGGCCAGGAGAGAG 507
Db 446 CTATGATCTTTTATATGCCATCTCCAAAGCTGGGGCCAGGAGAGAG 491

RESULT 15

US-09-724-676A-12749
; Sequence 12749, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12749
; LENGTH: 911
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature

```
; LOCATION: (11)..(11)
; OTHER INFORMATION: n is a,c,g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (18)..(18)
; OTHER INFORMATION: n is a,c,g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (23)..(23)
; OTHER INFORMATION: n is a,c,g, or t
; US-09-724-676A-12749
```

```
Query Match 56.8%; Score 340.4; DB 5; Length 911;
Best Local Similarity 84.5%; Pred. No. 8e-97;
Matches 394; Conservative 0; Mismatches 71; Indels 1; Gaps 1;
```

```
OY 43 GAGTCGTTGTTAGCGCTGCTGCCCCGTGAAAAAGCAAGATGATTTTAGGACTTTAGCTA 102
Db 26 GGTGGGCTGTGAGCGCTTTCAGTGAAGAAAGCAAGATGATTTTAGAACTTTAACTA 85
OY 103 CAATTTAGCTCCCTGATTCCTGTGTGCGCAACGATGATCTTCAAGGAAAGTAGAC 162
Db 86 CAGTTTAGCTCCCTGATTCCTGTGTGCGCAACGATGATCTTCAAGGAAAGTAGAAC 145
OY 163 CAGGCGCAATTGATCCCTCTCGAACATATGACAAAGACACCACTTCAGATTT 222
Db 146 CAGGCGCAATTGATCCCTCTCGAACATATGACAAAGACACCACTTCAGATTT 205
OY 223 TAAGAGCTTCAACGCTGTCGGATAACTTCAGCAACCCCTTATCTGACGCGGATGCCAG 282
Db 206 TAAGAGCTTCAACGCTGTCGGATAACTTCAGCAACCCCTTCTCGACGAGATGCCAG 265
OY 283 GCTGGCGCTGCAACAGACCGAGTTCCTGCGGAAGAAATGAAGTTGTATTTGCTCAGAC 342
Db 266 GCTCCAGCTGCATAGACTGAGTTTCTGGAAAAGAAATGAAGTTATTTGCTCAGAC 325
OY 343 TTACACATAGGAAGTTCAACCTGCTCGGC-AAATCCGACAAACAGTTCTCATCTC 401
Db 326 CTTACACATAGGAAGTTCAACCTGCTCGGC-AAATCCGACAAACAGTTCTCATCTC 385
OY 402 CCCTCCGCGCTCTCTCTCCGTTGCTGGAACAAGTAGAAGATGCCACCCCGTCATAAA 461
Db 386 CCCTCCGCGCTCTCTCTCCGTTGCTGGAACAAGTAGAAGATGCCACCCCGTCATAAA 445
OY 462 TTACGATCTTTATATGATCCATCCAAAGCTGCGGCGCAGAGAGAGAG 507
Db 446 CTTATGATCTTTATATGATCCATCCAAAGCTGCGGCGCAGAGAGAGAG 491
```

Search completed: December 14, 2002, 19:59:16
Job time : 37.2872 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 14, 2002, 21:32:16 ; Search time 33.0556 Seconds
(without alignments)
2368.771 Million cell updates/sec

Title: US-09-782-953-7
Perfect score: 1043
Sequence: 1 MDRDRSYNFSLLIACVAND.....RPKPKIIQTRRPEYTPHLS 198

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 355320 seqs, 197730502 residues

Total number of hits satisfying chosen parameters: 710640

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DRV=xlh
-Q=/cgn2_1/USPTO.spool/US09782953/runat_11122002_114431_17465/app.query.fasta_1.1173
-DB=Published Applications NA -QWMT=fastcap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09782953 -CGN 1 1 38 @runat_11122002_114431_17465
-NCFU=6 -ICPU=3 -NO_XLPAY -NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/FCR_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1043	100.0	597	10	US-09-782-953-5
2	970.5	93.0	2331	10	Sequence 5, Appli
3	970.5	93.0	2355	12	Sequence 11, Appl
4	920.5	88.3	597	10	Sequence 255, App
					Sequence 2, Appli

5	875	83.9	2212	10	US-09-782-953-14	Sequence 14, Appli
6	868	83.2	599	10	US-09-782-953-1	Sequence 1, Appli
7	847.5	81.3	2173	10	US-09-880-107-3340	Sequence 3340, Ap
8	830	79.6	2358	10	US-09-925-102-347	Sequence 347, App
9	632.5	60.6	594	10	US-09-782-953-8	Sequence 8, Appli
10	615.5	59.0	3184	10	US-09-954-456-497	Sequence 497, App
11	615.5	59.0	3184	10	US-09-782-953-17	Sequence 17, Appl
12	592.5	56.8	828	10	US-09-782-953-20	Sequence 20, Appl
13	591.5	56.7	720	10	US-09-782-953-23	Sequence 23, Appl
14	285.5	27.4	446	10	US-09-864-761-10388	Sequence 10388, A
15	285.5	27.4	446	10	US-09-864-761-2064	Sequence 2064, Ap
16	266.5	25.6	486	10	US-09-864-761-809	Sequence 809, App
17	200	19.2	365	10	US-09-728-445-736	Sequence 736, App
18	187.5	18.0	123	10	US-09-864-761-17592	Sequence 17592, A
19	181	17.4	111	10	US-09-864-761-17592	Sequence 17592, A
20	140	13.4	85	10	US-09-864-761-17592	Sequence 17592, A
21	90.5	8.7	12308	9	US-09-854-133-422	Sequence 18808, A
22	90.5	8.7	12308	10	US-09-738-973-422	Sequence 422, App
23	87	8.3	4600	9	US-09-738-973-422	Sequence 422, App
24	87	8.3	4600	9	US-09-902-941-1797	Sequence 1797, Ap
25	82.5	7.9	530	10	US-09-864-761-15897	Sequence 15897, A
26	81.5	7.8	2095	12	US-10-044-090-665	Sequence 665, App
27	81	7.8	3138	10	US-09-880-107-1716	Sequence 1716, App
28	81	7.8	7177	12	US-10-033-026-7	Sequence 7, Appli
29	81	7.8	7364	10	US-09-954-456-1179	Sequence 1179, Ap
30	81	7.8	7364	12	US-10-033-026-5	Sequence 5, Appli
31	81	7.8	7376	12	US-10-033-026-3	Sequence 3, Appli
32	81	7.8	640681	10	US-09-790-988-1	Sequence 1, Appli
33	80.5	7.7	6799	9	US-09-902-941-1883	Sequence 1883, Ap
34	80	7.7	552	10	US-09-954-456-594	Sequence 594, App
35	79	7.6	3775	10	US-09-954-456-2210	Sequence 2210, Ap
36	79	7.6	3775	10	US-09-880-107-2221	Sequence 2221, Ap
37	79	7.6	36651	10	US-09-964-469-3	Sequence 3, Appli
38	78.5	7.5	569	9	US-10-040-739-301	Sequence 301, App
39	77.5	7.4	4108	10	US-09-883-096-1	Sequence 1, Appli
40	77.5	7.4	32191	10	US-09-764-877-3374	Sequence 3374, Ap
41	77	7.4	1440	9	US-09-712-363-94	Sequence 94, Appl
42	77	7.4	2149	10	US-09-880-107-3031	Sequence 3031, Ap
43	77	7.4	5739	10	US-09-960-253-142	Sequence 142, App
44	77	7.4	6417	10	US-09-962-436-288	Sequence 288, App
45	76.5	7.3	780	9	US-09-938-842A-1412	Sequence 1412, Ap

ALIGNMENTS

RESULT 1
US-09-782-953-5
; Sequence 5, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: US/09-782-953
; CURRENT APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 5
; LENGTH: 597
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(594)
US-09-782-953-5
Alignment Scores: 1.33e-118 Length: 597
Pred. No.: 1043.00 Matches: 198

Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 100.00%
 DB: 10
 Gaps: 0

US-09-782-953-7 (1-198) x US-09-782-953-5 (1-597)

QY 1 MetAspPheArgAspPheSerTyraenPheSerSerLeuIleAlaCyValAlaAsnAsp 20
 DB 1 ATGATTTTGGGACTTTTACCTAATTTTACCTCCGATTTGCTGTGGCAACAT 60
 QY 21 AspValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThrTyraP 40
 DB 61 GATGCTTCAGCGAAAGTAGACACAGGGCCAAATTTGAATCCCTCTTCAGAACATATGAC 120
 QY 41 LysAspThrThrPheGlnTyrrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60
 DB 121 AAGGACACACACCTTCAGATTTTAAAGACTTCMAACGTCCGGATTAACCTTCAGCAAC 180
 QY 61 ProLeuSerAlaAlaAspAlaArgLeuAlaLysLeuAlaLysThrGluPheLeuGluLysGlu 80
 DB 181 CCTTATCTGCAGACCCGATCCAGGCTGGCGTGCACAAAGACCGAGTTCTGGGAAAGGA 240
 QY 81 MetLysLeuTyrrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProProAsn 100
 DB 241 ATGAAGTTATTTTGTCTCAGACTTTACACATAGAAAGTTTCACTGCTCCGCCCAAT 300
 QY 101 ProAspLysGlnPheLeuLysSerProProAlaSerProProValGluTyrrLysGlnVal 120
 DB 301 CCGGACAAACAGTCTCATCTCCCTCCGCGCTTCTCTCCCTGGCTGGGAAACAGTA 360
 QY 121 GluAspAlaThrProValIleAsnTyrrAspLeuLeuTyrrAlaIleSerLysLeuGlyPro 140
 DB 361 GAGAGTCCACCCCGCATMAATTCGATCTTTATATGCCATCTCCAAAGCTGGGGCCA 420
 QY 141 GlyGluLysTyrrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160
 DB 421 GGAGAGAAGATGATGATGCTGACGAGACAGCCCACTCCAGTGTGGTCCAGCTG 480
 QY 161 CysGluSerAspGlnGluAsnGluGluGluGluGluGluGluGluGluGluGluGluGlu 180
 DB 481 TGTGAGAGTACCAAGATGAGATGAGAGAGAGAGAGAGAGAGATGAGAGAGAGAGAGAGAG 540
 QY 181 LysProLysIleIleGlnThrArgArgProGluTyrrThrProIleHisLeuSer 198
 DB 541 AAGCCAAATATCTCCAGACACGAGACCGAGATACACAGATCCACCTTAC 594

RESULT 2

US-09-782-953-11
 Sequence 11, Application US/09782953

Patent No. US20020150953A1
 GENERAL INFORMATION:

APPLICANT: WILLIAMS, R. SANDERS
 APPLICANT: ROTHERMEL, BEVERLY

TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
 FILE REFERENCE: US/09-782-953-11

CURRENT APPLICATION NUMBER: US/09-782-953
 PRIOR FILING DATE: 2001-02-13

PRIOR APPLICATION NUMBER: 60/216,601
 NUMBER OF SEQ ID NOS: 27

SOFTWARE: Patentin Ver. 2.1
 SEQ ID NO 11

LENGTH: 2331
 TYPE: DNA

ORGANISM: Homo sapiens
 FEATURE:

NAME/KEY: CDS
 LOCATION: (144) .. (734)

US-09-782-953-11

Alignment Scores:

Pred. No.: 7.13e-109
 Score: 970.50
 Percent Similarity: 96.97%
 Best Local Similarity: 93.43%
 Query Match: 93.05%
 DB: 10
 Gaps: 1

US-09-782-953-7 (1-198) x US-09-782-953-11 (1-2331)

QY 1 MetAspPheArgAspPheSerTyraenPheSerSerLeuIleAlaCyValAlaAsnAsp 20
 DB 144 ATGATTTTGGGAACTTTTACCTAATTTTACCTCCGATTTGCTGTGGCAACAT 203
 QY 21 AspValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThrTyraP 40
 DB 204 GATATCTTCAGCGAAAGTAGACACAGGGCCAAATTTGAATCCCTCTTCAGAACATATGAC 263
 QY 41 LysAspThrThrPheGlnTyrrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60
 DB 264 AAGGACACACACCTTCAGATTTTAAAGACTTCMAACGTCCGGATTAACCTTCAGCAAC 323
 QY 61 ProLeuSerAlaAlaAspAlaArgLeuAlaLysLeuAlaLysThrGluPheLeuGluLysGlu 80
 DB 324 CCTTATCTGCAGACCCGATCCAGGCTGGCGTGCACAAAGACCGAGTTCTGGGAAAGGA 383
 QY 81 MetLysLeuTyrrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProProAsn 100
 DB 384 ATGAAGTTATTTTGTCTCAGACTTTACACATAGAAAGTTTCACTGCTCCGCCCAAT 443
 QY 101 ProAspLysGlnPheLeuLysSerProProAlaSerProProValGluTyrrLysGlnVal 120
 DB 444 CCAAGCAACAGTTCCTGATTCCTCCCTCCGCGCTTCTCTCCCTGGCTGGGAAACAGTA 503
 QY 121 GluAspAlaThrProValIleAsnTyrrAspLeuLeuTyrrAlaIleSerLysLeuGlyPro 140
 DB 504 GAAATGCGACCCCGCATMAATTCGATCTTTATATGCCATCTCCAAAGCTGGGGCCA 563
 QY 141 GlyGluLysTyrrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160
 DB 564 GGGGAAAGTATGAAATGACGAGACAGCTACACCACTCCAGCTGGTGTGCTCATGTA 623
 QY 161 CysGluSerAspGlnGluAsnGluGluGluGluGluGluGluGluGluGluGluGluGlu 180
 DB 624 TGTGAGAGTATCAAGAG---AAGGAGAGAGAGAGAGAGAGATGAGAGAGAGAGAGAGAG 680
 QY 181 LysProLysIleIleGlnThrArgArgProGluTyrrThrProIleHisLeuSer 198
 DB 681 AAGCCAAATATCTCCAGACACGAGACCGAGATACACAGATCCACCTTAC 734

RESULT 3

US-10-044-090-255

Sequence 255, Application US/10044090
 Patent No. US20020137081A1

GENERAL INFORMATION:

APPLICANT: Olga Bandman
 TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION

FILE REFERENCE: PA-0028 US
 CURRENT APPLICATION NUMBER: US/10/044,090

CURRENT FILING DATE: 2002-01-09
 NUMBER OF SEQ ID NOS: 850

SOFTWARE: PERL Program
 SEQ ID NO 255

LENGTH: 2355
 TYPE: DNA

ORGANISM: Homo sapiens
 FEATURE:

NAME/KEY: misc feature
 OTHER INFORMATION: Incyte ID NO. US20020137081A1 5511889CB1

US-10-044-090-255

Alignment Scores: 7.24e-109 Length: 2355
 Pred. No.: 970.50 Matches: 185
 Score: 970.50

Percent Similarity: 96.97% Conservative: 7
 Best Local Similarity: 93.43% Mismatches: 5
 Query Match: 93.05% Indels: 1
 DB: 12 Gaps: 1

US-09-782-953-7 (1-198) x US-10-044-090-255 (1-2355)

```

Qy 1 MetAspPheArgAspPheSerTyrAsnPheSerSerLeuLeuAlaCysValAlaAsnAsp 20
Db 151 ATGCATTTTGAAGAACTTTAACTACAGTTTATAGCTCCCTGATTGCTGTGTGGCAACAGT 210
Qy 21 AspValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40
Db 211 GATATCTTCAGCGAAAGTGAACCAAGGCCAAATTTGAGTCCCTCTTTAGGACGATGAC 270
Qy 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60
Db 271 AAGGACATCACCTTTCAGTATTTTAAAGAGCTTCAACAGAGTTCAGAAATAAATTCAGCAAC 330
Qy 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
Db 331 CCTTCTCCGAGCAGATGCCAGCTCCAGCTGCATAAGACTGAGTTTCTGGGAAAGGAA 390
Qy 81 MetLysLeuTyrPheAlaGlnThrLeuHisLysGlySerSerHisLeuAlaProProAsn 100
Db 391 ATGAAGTTATATTGCTCAGACCTTACATAGGAAGCTCACACCTGGCTCCGCCAAAT 450
Qy 101 ProAspLysGlnPheLeuLeuSerProProAlaSerProValGlyTyrLysGlnVal 120
Db 451 CCAGACAAGCAGTTTCTGATCTCCCTCCCGCTCTCCGCGAGTGGAGTGAACAAGTG 510
Qy 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaLysSerLysLeuGlyPro 140
Db 511 GAAGATCGGACCCAGTCATAAATATGATCTCTTATATGCCATCTCCAAAGCTGGGGCCA 570
Qy 141 GlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160
Db 571 GGGGAAAAGTATGAATTCAGCAGCAGCAGCTGACACCACTCCAGCGTGGTGGTCCATGTA 630
Qy 161 CysGluSerAspGlnGluAsnGluGluGluGluMetGluArgMetLysArgPro 180
Db 631 TGTGAGAGTATCAAGAG---AAGGAGGAAGAGAGGAAATGGAAGATGAGGAGACCT 687
Qy 181 LysProLysIleLeuGlnThrArgAtqProGluTyrThrProLysLeuSer 198
Db 688 AAGCCAAAATATTCAGACAGGAGCGCGAGTACACGCCGATCCACCTCAGC 741

```

RESULT 4

```

US-09-782-953-2
; Sequence 2, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: US/09782953
; CURRENT APPLICATION NUMBER: US/09782953
; PRIOR FILING DATE: 2001-02-13
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 597
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(594)
US-09-782-953-2

```

Alignment Scores:

Pred. No.: 1,23e-103 Length: 597
 Score: 920.50 Matches: 178
 Percent Similarity: 92.42% Conservative: 5
 Best Local Similarity: 89.90% Mismatches: 12
 Query Match: 88.26% Indels: 3
 DB: 10 Gaps: 1

US-09-782-953-7 (1-198) x US-09-782-953-2 (1-597)

```

Qy 1 MetAspPheArgAspPheSerTyrAsnPheSerSerLeuLeuAlaCysValAlaAsnAsp 20
Db 10 GTGGATCTCGAGGACCTGCGC-----AGCGCCACCATCGCTGCCACCTGGACCCG 60
Qy 21 AspValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40
Db 61 CCGTGTTCGTGGAGCGCTGTGCGGGCCAAATTTGAATCCCTCTTCAGAACATATGAC 120
Qy 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60
Db 121 AAGGACACCATCTCCAGTATTTTAAAGAGCTTCAACAGTGTCCGGATAAATTCAGCAAC 180
Qy 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
Db 181 CCTTATCTGCGAGCGATGCCAGGCTGCGGTGCAACAAGACCGAGTTCTTGGGAAGGAA 240
Qy 81 MetLysLeuTyrPheAlaGlnThrLeuHisLysGlySerSerHisLeuAlaProProAsn 100
Db 241 ATGAAGTTGATTTTGTCTCAGACTTTACATAGGAAGTTCACACCTGGCTCCGCCAAAT 300
Qy 101 ProAspLysGlnPheLeuLeuSerProProAlaSerProValGlyTyrLysGlnVal 120
Db 301 CCGGACAAACAGTCTCTCATCTCCCTCCGGCTCTCTCCGTTGGTGGTGAACAAGTA 360
Qy 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaLysSerLysLeuGlyPro 140
Db 361 GAAGATGCCACCCCGCTCATAAATTAACGATCTTTATATGCCATCTCCAAAGCTGGGGCCA 420
Qy 141 GlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160
Db 421 GGAGAGAAGTATGAATTCATGTCAGCGACAGACCCACTCCAGTGTGTGTGTCCACGTG 480
Qy 161 CysGluSerAspGlnGluAsnGluGluGluGluMetGluArgMetLysArgPro 180
Db 481 TGTGAGAGTGAACCAAGAGATGAGGAGGAGAGAGAGATGAGAGAGATGAGAGAGACCC 540
Qy 181 LysProLysIleLeuGlnThrArgAtqProGluTyrThrProLysLeuSer 198
Db 541 AAGCCAAAATATTCAGACAGGAGCGCGAGTACACCGATCCACCTTAGC 594

```

RESULT 5

```

US-09-782-953-14
; Sequence 14, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: US/09782953
; CURRENT APPLICATION NUMBER: US/09782953
; PRIOR FILING DATE: 2001-02-13
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 2212
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (25)..(615)
US-09-782-953-14

```

Alignment Scores:
 Pred. No.: 3,08e-97 Length: 2212
 Score: 875.00 Matches: 171
 Percent Similarity: 90.40% Conservative: 8
 Best Local Similarity: 86.36% Mismatches: 15
 Query Match: 83.89% Indels: 4
 DB: 10 Gaps: 2

US-09-782-953-7 (1-198) x US-09-782-953-14 (1-2212)

QY 1 MetAapPheArGApPheSerTyRanPheSerSerLeuIleAlaCyValAlaAnaap 20
 DB 34 GTGGAGCTCGACAGACCTGCCCC-----AGCGCCACCATGCTGTTCACCTGACCG 84
 QY 21 AapValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThrTyRasp 40
 DB 85 CCGGTGTCTGTGACCGCTGTGCGGCCAAATTGAGTCCCTCTTAAAGACGATGAC 144
 QY 41 LysAapThrThPheGlnTyRPhelySerPheLysArgValArgIleAnPheSerAn 60
 DB 145 AAGGACATCACCTTTCAGATTTTAAAGACTTCAACAGCTCAGAAATAACTTCAGCAAC 204
 QY 61 ProLeuSerAlaAlaAapAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
 DB 205 CCTTCTCCGACAGATGCGACGCTCCAGCTGCATAGACTGAGTTTCTGGGAAAGGAA 264
 QY 81 MetLysLeuTyRPhelAglnThrLeuHisIleGlySerSerHisLeuAlaProProAn 100
 DB 265 ATGAGTATATATTTCTCTGACCTTACACATAGAAAGCTCACACCTGCTCCGCAAT 324
 QY 101 ProAapLysGlnPheLeuLysSerProProAlaSerProProValGlyTpyLysGlnVal 120
 DB 325 CCGACACAGAGATTTCTGATCTCCCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCT 384
 QY 121 GluAapAlaThrProValIleAenTyRaspLeuLeuTyRAlaIleSerTyLysGlyPro 140
 DB 385 GAGAGTCCGACCCAGCATATACTATGATCTTATATCCATCTTCCAAAGCTGGGGCCA 444
 QY 141 GlyGluLysTyRGlLeuHisAlaAlaThrAapProThrProSerValValIleHisVal 160
 DB 445 GGGGAAAGATGAAATTCAGACGACGACGACGACGACGACGACGACGACGACGACGAC 504
 QY 161 CysGluSerAapGlnGluAenGluGluGluGluGluGluGluGluGluGluGluGlu 180
 DB 505 TGTGAGCTGATCAAGAG---AGGAGGAAAGAGAGGAATGGAAGAAATGAGAGACCT 561
 QY 181 LysProLysIleIleGlnThrArgArgProGluTyRThrProIleHisLeuSer 198
 DB 562 AAGCAAAATTTATCCAGACCGAGGCGGAGTACAGCCGATCCACCTCAGC 615
 RESULT 6
 US-09-782-953-1
 ; Sequence 1, Application US/09782953
 ; Patient No. US20020150953A1
 ; GENERAL INFORMATION:
 ; APPLICANT: WILLIAMS, R. SANDERS
 ; APPLICANT: ROTHMEL, BEVERLY
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
 ; FILE REFERENCE: UTS:674P21
 ; CURRENT APPLICATION NUMBER: US/09/782,953
 ; PRIOR FILING DATE: 2001-02-13
 ; PRIOR APPLICATION NUMBER: 60/216,601
 ; PRIOR FILING DATE: 2000-07-07
 ; NUMBER OF SEQ ID NOS: 27
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 599
 ; TYPE: DNA
 ; ORGANISM: Mus musculus
 US-09-782-953-1

Alignment Scores:
 Pred. No.: 3.22e-97 Length: 599
 Score: 868.00 Matches: 170
 Percent Similarity: 98.84% Conservative: 0
 Best Local Similarity: 98.84% Mismatches: 2
 Query Match: 83.22% Indels: 1
 DB: 10 Gaps: 0

US-09-782-953-7 (1-198) x US-09-782-953-1 (1-599)

QY 2 AapPheArGApPheSerTyRanPheSerSerLeuIleAlaCyValAlaAnaap 21
 DB 83 GATTTTGAGACTTTAGCTTCATTTTAACTTCTGATGCTGTGGGAAAGATGAT 142
 QY 22 ValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThrTyRasp 41
 DB 143 GTCTTCAGCAAGTGTAGACAGGCGCAATTGAAATCCCTTCAGAACATATGACAA 202
 QY 42 AapThrThPheGlnTyRPhelySerPheLysArgValArgIleAnPheSerAnPro 61
 DB 203 GACACACCTTCCAGATTTTAAAGAGCTTCAACGATGTCGGAATAAATTCAAGCAACCC 262
 QY 62 LeuSerAlaAlaAapAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGluMet 81
 DB 263 TTATCTGACGCCATTCGACGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCT 322
 QY 82 LysLeuTyRPhelAglnThrLeuHisIleGlySerSerHisLeuAlaProProAnPro 101
 DB 323 AAGTGTATTTGCTCAGACTTACATAGAAAGTTACACCTGCTCCGCTCCGCTCCGCT 381
 QY 102 AapLysGlnPheLeuLysSerProProAlaSerProProValGlyTpyLysGlnVal 121
 DB 382 GACAAACAGTTCATCTCCCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCT 441
 QY 122 AapAlaThrProValIleAenTyRaspLeuLeuTyRAlaIleSerTyLysGlyProGly 141
 DB 442 GATGCAACCCCGATCATATAATTAGATCTTTATATGCAATCTCAGAGTGGGCCAGGA 501
 QY 142 GlyLysTyRGlLeuHisAlaAlaThrAapProThrProSerValValIleHisVal 161
 DB 502 GAGAGTATGAACTGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 561
 QY 162 GluSerAapGlnGluAenGluGluGluGluGluGluGluGluGluGluGluGlu 173
 DB 562 GAGAGTGAACCAAGAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 597
 RESULT 7
 US-09-880-107-3340
 ; Sequence 3340, Application US/09880107
 ; Patient No. US20020142981A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Horne, Darci T.
 ; APPLICANT: Vockley, Joseph G.
 ; APPLICANT: Scherf, Uwe
 ; APPLICANT: Gene Logic, Inc.
 ; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
 ; FILE REFERENCE: 44921-5028-MO
 ; CURRENT APPLICATION NUMBER: US/09/880,107
 ; PRIOR FILING DATE: 2001-06-14
 ; PRIOR APPLICATION NUMBER: US 60/211,379
 ; PRIOR FILING DATE: 2000-06-14
 ; PRIOR APPLICATION NUMBER: US 60/237,054
 ; PRIOR FILING DATE: 2000-10-02
 ; NUMBER OF SEQ ID NOS: 3950
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 3340
 ; LENGTH: 2173
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: Genbank Accession No. US20020142981A1 U28833
 US-09-880-107-3340

Alignment Scores:

Pred. No.: 6.87e-94 Length: 2173
 Score: 847.50 Matches: 162
 Percent Similarity: 97.63% Conservative: 3
 Best Local Similarity: 95.86% Mismatches: 3
 Query Match: 81.26% Indels: 1
 DB: 10 Gaps: 1

US-09-782-953-7 (1-198) x US-09-880-107-3340 (1-2173)

QY 30 AlaLysPheGluSerLeuPheArgThrTyrAspLysAspThrPheGlnTyrPheLys 49
 DB 58 GCCAAATTTAGTCCCTTTTAGACGTATGACAGGACATCACCTTTTCAGTATTTTAAAG 117
 QY 50 SerPheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAspAlaArgLeu 69
 DB 118 AGCTTCAACAGGTGAGTAACTTCAGCAACCCCTTCTCCGACGAGATGCCAGGCTC 177
 QY 70 ArgLeuHisLysThrGluPheGluGlyLysGluMetLysLeuTyrPheAlaGlnThrLeu 89
 DB 178 CAGCTGCATAAGACTGAGTTTCTGGGAAGGAAATGAAGTTATATTTTGTCTCAGACCTTA 237
 QY 90 HisIleGlySerSerHisLeuAlaProProAsnProAspLysGlnPheLeuIleSerPro 109
 DB 238 CACATAGGAGCTCACACCTGGCTCCGCCAAATCCACAGCAGGTTCCTGATCTCCCT 297
 QY 110 ProLysProProValGlyTTrpLysGlnValGluAspAlaThrProValIleAsnTyr 129
 DB 298 CCCGCCCTCTCCGCGTGGATGAAACAAGTGAAGATCGGACCCAGTCATAACTAT 357
 QY 130 AspLeuLeuTyrAlaIleSerLysLeuGlyProGlyGluLysTyrGluLeuHisAlaAla 149
 DB 358 GATCTCTTATATGATCCATCTCCAGCTGGGCGCAGGGAAGATGAATTTGCACGCGAGCG 417
 QY 150 ThrAspProThrProSerValValHisValCysGluSerAspGlnGluAsnGluGlu 169
 DB 418 ACTGACACCACTCCACCGTGGTCCATGTATGTGAGAGTATCAAGAG---AAGGAG 474
 QY 170 GluGluGluGluMetGluArgMetLysArgProLysProLysIleIleGlnThrArgArg 189
 DB 475 GAAGAAGAGGAAATGGAAGAATGAGGAGACCTAAGCCAAAATTTATCCAGACGAGGAG 534
 QY 190 ProGluTyrThrProIleHisLeuSer 198
 DB 535 CCGGAGTACACGCGATCCACCTCAGC 561

RESULT 8

US-09-925-302-347
 ; Sequence 347, Application US/09925302
 ; Patent No. US20020044941A1

GENERAL INFORMATION:

; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
 ; FILE REFERENCE: PA104
 ; CURRENT APPLICATION NUMBER: US/09/925,302
 ; PRIOR FILING DATE: 2001-08-10
 ; PRIOR APPLICATION NUMBER: PCT/US00/05918
 ; PRIOR FILING DATE: 2000-03-08
 ; PRIOR APPLICATION NUMBER: 60/124,270
 ; PRIOR FILING DATE: 1999-03-12
 ; NUMBER OF SEQ ID NOS: 896
 ; SOFTWARE: Patentin Ver. 2.0
 ; SEQ ID NO 347
 ; LENGTH: 2358
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens

US-09-925-302-347

Alignment Scores:

Pred. No.: 1.06e-91 Length: 2358
 Score: 830.00 Matches: 168
 Percent Similarity: 88.89% Conservative: 8
 Best Local Similarity: 84.85% Mismatches: 18

Query Match: 79.58% Indels: 5
 DB: 10 Gaps: 2
 US-09-782-953-7 (1-198) x US-09-925-302-347 (1-2358)
 QY 1 MetAspPheArgAspPheSerTyrAsnPheSerSerLeuIleAlaCysValAlaAsnAsp 20
 DB 123 GTGGACCTGCAGGACCTGCC-----AGGCCACCATGCCTGTCACTGGACCGG 173
 QY 21 AspValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40
 DB 174 CGCGTGTTCGTGGACGCGCTGTGCGGGCCAAATTTAGTCCCTCTTTAGGAGGTATGAC 233
 QY 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60
 DB 234 AAGGACATCACCTTTTCAGTATTTTAAGAGCTTCAACAGAGTCAGATAAATTCAGCAAC 293
 QY 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
 DB 294 CCCTTCTCCGACGAGATGCCAGGCTCCAGCTGCATAGACTGAGTTTCTGGGAAGGAA 353
 QY 81 MetLysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProProAsn 100
 DB 354 ATGAAGTTATATTTTGTCTCAGACCTTACATAGGAAGCTCACACTGGCTCCGCA-AA 412
 QY 101 ProAspLysGlnPheLeuIleSerProProAlaSerProProValGlyTrpLysGlnVal 120
 DB 413 CCAGACAGCAGATTTCTGATCTCCCTCCCGCTCTCCGSCAGTGGGATGGAACAAGTG 472
 QY 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro 140
 DB 473 GAAGATGGACCCAGTCATAACTATGATCTTTATATGCCATCTCCAGCTGGGCA 532
 QY 141 GlyLysTyrGluLeuHisAlaThrAspProThrProSerValValHisVal 160
 DB 533 GGGGAAAGATGAATTCAGCAGCAGCTGACACCACTCCAGCGTGGTGTCCATGTA 592
 QY 161 CysGluSerAspGlnGluAsnGluGluGluGluMetGluArgMetLysArgPro 180
 DB 593 TGTAGAGTGTCAAGAG---AAGGAGGAAGAGGAATGGAAGAATGAGGAGACCT 649
 QY 181 LysProLysIleIleGlnThrArgArgProGluTyrThrProIleHisLeuSer 198
 DB 650 AAGCCAAAATTTATCCAGACGAGGCGGAGTACACGCGATCCACCTCAGC 703

RESULT 9

US-09-782-953-8

; Sequence 8, Application US/09782953
 ; Patent No. US20020150953A1

GENERAL INFORMATION:

; APPLICANT: WILLIAMS, R. SANDERS
 ; APPLICANT: ROTHERMEL, BEVERLY
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
 ; FILE REFERENCE: UTSID:674P21
 ; CURRENT APPLICATION NUMBER: US/09/782,953
 ; CURRENT FILING DATE: 2001-02-13
 ; PRIOR APPLICATION NUMBER: 60/216,601
 ; PRIOR FILING DATE: 2000-07-07
 ; NUMBER OF SEQ ID NOS: 27
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 8
 ; LENGTH: 594
 ; TYPE: DNA
 ; ORGANISM: Mus musculus
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(591)

US-09-782-953-8

Alignment Scores:

Pred. No.: 1.89e-68 Length: 594
 Score: 632.50 Matches: 121

```
Percent Similarity: 77.13%
Best Local Similarity: 64.36%
Query Match: 60.64%
DB: 10
Conservative: 2
Mismatches: 3
Indels: 9
Gaps: 2
```

US-09-782-953-7 (1-198) X US-09-782-953-8 (1-594)

```

Oy 11 SerSerLeuIIeAlaICyValAlaIaenAapAValIPheSerGIuSerGIuThrArxAla 30
Db 31 TCCACTCTGGTGGCCCTGTGTGGATGTGGAGGCTCTTACCATCAGAGACGTTAAAGAA 90
Oy 31 LysPheGIuSerLeuPheArgThrTyraSPuYaaPThrTrhPheGlnTyPheYsser 50
Db 91 AAATTGAGGAGCACTGTTCCGGAGCACTTATGATGAATGTGTGACGTTCCAGCTGTTAAAGAGT 150
Oy 51 PheLYaaYValaIArgIIeaenPheSerAenProLeuSerAlaIaAPAlaArgLeuArg 70
Db 151 TTCGCGCGGGTTCGATAAAATTTCAGCCATCCCAATCTGCAGCCCGCGGCCGGATAGAG 210
Oy 71 LeuHIaLYeThrGIuPheLeuGIuLYeGIuMeLysLeuTyPheAlaGlnThrLeuHis 90
Db 211 CTTTCATGACACTCAGTTTCAGAGGAAAGAAAGCTAAAACTCTAATTGCGCCAGGTCCAGACC 270
Oy 91 IlegYSerSer-----HisLeuAlaProProAenProAapPheGIuPheLeu 106
Db 271 CCAGAGACAGATGAGAGACAAACTGCATTTGGGACCTCCAGCCAGCCGCAACAGATCTTC 330
Oy 107 ILeSerProProAlaSerProProValGIYTPuLYeGlnValaIGluAaPAlaThrProVal 126
Db 331 ATCTACCCCTTCATCTCTCTCTCTGTGGCTGGAACTTATGAGGAGATGCCACACATC 390
Oy 127 ILeaenTyraSPuLeuLeuTyraIaIeSerLYeLeuGIYProGIuLYeTyGIuLeu 146
Db 391 CTCACATGACCTTCTTATGCTGTGGCCAAACTAGGACCGAGAGAGAAATATAGTGTG 450
Oy 147 HIaAlaAlaThrAapProThrProSerValIValaIaValaCYeGIuSerAapGlnGIu 166
Db 451 CAGCTGGAACCTGAGCTACACCGAGCGTCTGTGTGATCTGTGTGACAGCAGCATGAGAG 510
Oy 167 AenGIuGIuGIuGIuGIuMeGIuArgMetuYaaArgProLYaSPuLYaIleIleGln 186
Db 511 GAGAGAGAGAGACCCAAAG-----ACTTCCCAAGCAGCAAAATCATTCAG 555
Oy 187 ThrArgArgProGIuTyThrPro 194
Db 556 ACCCGGCGTCCGGGCTTCCACCC 579

RESULT 10
US-09-954-456-497
Sequence 497, Application US/09954456
Patent No. US20020115057A1
GENERAL INFORMATION:
APPLICANT: Young, Paul
TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc
FILE REFERENCE: 689290-76
CURRENT APPLICATION NUMBER: US/09/954,456
CURRENT FILING DATE: 2001-09-18
PRIOR APPLICATION NUMBER: US/60/233,617
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: US/60/234,052
PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: US/60/234,923
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US/60/235,134
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US/60/235,637
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: US/60/235,638
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: US/60/235,711
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,720

```

```

; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/225,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/225,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 497
; LENGTH: 3184
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-497

Alignment Scores:
Pred. No.: 2,686-65 Length: 3184
Score: 615.50 Matches: 119
Percent Similarity: 75.53% Conservative: 23
Best Local Similarity: 63.30% Mismatches: 37
Query Match: 59.01% Indels: 9
DB: 10 Gaps: 2

US-09-782-953-7 (1-198) x US-09-954-456-497 (1-3184)
QY 11 SerSerLeuLeuAlaCyValAlaAsnAspAspAlaPheSerGluSerGluThrArgAla 30
Db 220 TCCACTCGTGTTGCTGATGTGTGGATGATCGAGGCTTACCAATCAGAGTTAAAGAA 279
QY 31 LysPheGluSerLeuPheArgThrTyrArgPheAspThrThrPheGlnTyrPheLysSer 50
Db 280 AAATTGGGGGAGCTGTTTCGGACTTAATGATACGTGTGAGGTTCCAGCTTAATTAAGAT 339
QY 51 PheLysArgValArgIleAsnPheSerAsnProLeuSerAlaIleAspAlaArgLeuArg 70
Db 340 TTCAGACGTGTCGGATTAATCTTCAGCAATCTTAATCTGCAGCCCGCAGCTAGATAGAG 399
QY 71 LeuHISLysThrGluPheLeuGlyLysGluMetLysLeuTyrPheAlaGlnThrLeuHIS 90
Db 400 CTTCAATGAAGAACCAATTCAGAGGAAATAATTAAGCTCTACTTTCACAGATTACAGACT 459
QY 91 IleGlySerSer-----HisLeuAlaProProAsnProAspLysGlnPheLeu 106
Db 460 CCAGAGACAGATGAGAGACAAACTGCACCTTGCTCCACCCAGCCTGCCAAACAGTTTCTC 519
QY 107 IleSerProProAlaSerProProValGlyTrpLysGlnIleGluAspAlaThrProVal 126
Db 520 ATTCGCCCCCTTCCTCCACCTGTTAGCTGGACGCCCATTCAGATGCCACCCAGTC 579
QY 127 IleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyProGlyGluLysTyrGluLeu 146
Db 580 CTCACATTAAGACTCCTCTATGCTGTGGCAAACTAGAGCCAGGAGAGATATAGCTTC 639
QY 147 HisAlaAlaThrAspProThrProSerValValValHisValCysGluSerAspGlnIle 166
Db 640 CATGACGAGGACTGACGCCACCCCAAGTGCCTGCTGCACAGTGTGGACAGGACATTAAG 699
QY 167 AsnGluGluGluGluGluGluMetGluArgMetLysArgProLysProLysIleIleGln 186
Db 700 GAGGAAGAGAGACCAAG-----ACTTCCCAAGCAAAACCAAAATCATTCGAA 744
QY 187 ThrArgArgProGluTyrThrPro 194
Db 745 ACTCGGCGTCTGGCTGCACACC 768

RESULT 11
US-09-782-953-17
; Sequence 17, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; TITLE OF INVENTION: CALCIUM ION INTERACTING PROTEIN (MCIP)
; FILE REFERENCE: UTSD:674PZ1

```

; CURRENT APPLICATION NUMBER: US/09/782,953
 ; CURRENT FILING DATE: 2001-02-13
 ; PRIOR APPLICATION NUMBER: 60/216,601
 ; PRIOR FILING DATE: 2000-07-07
 ; NUMBER OF SEQ ID NOS: 27
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 17
 ; LENGTH: 3184
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (205)..(780)
 US-09-782-953-17

Alignment Scores:

Pred. No.: 2,686-65 Length: 3184
 Score: 615.50 Matches: 119
 Percent Similarity: 75.53% Conservative: 23
 Best Local Similarity: 63.30% Mismatches: 37
 Query Match: 59.01% Indels: 9
 DB: 10 Gaps: 2

US-09-782-953-7 (1-198) x US-09-782-953-17 (1-3184)

QY 11 SerSerLeuIleAlaCysValAlaAsnAspValPheSerGluSerGluThrArgAla 30
 DB 220 TCCACTCTGGTGGCTGTGGTGGATCTCGAGGTCTTTACCAATCAGAGGTTAAGGAA 279
 QY 31 LysPheGluSerLeuPheArgThrTyrAspLysAspThrPheGlnTyrPheLysSer 50
 DB 280 AAATTGGGGGACTGTTTCGACTATGATGACTGTGTGAGCTTCCAGCTATTAAAGT 339
 QY 51 PheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAspAlaArgLeuArg 70
 DB 340 TTCAGAGCTGTCGGTATAAACTTCAGCAATCTTAAATCTGCAGCCGAGCTAGATAGAG 399
 QY 71 LeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyrPheAlaGlnThrLeuHis 90
 DB 400 CTTCAAGAACCAATTCAGAGGGAATAAATTAAGCTCTACTTTGCACAGGTTTCAGACT 459
 QY 91 IleGlySerSer-----HisLeuAlaProProAsnProAspLysGlnPheLeu 106
 DB 460 CCAGAGACAGATGAGACAACTGCACTTGGCTCCACCCCGAGCTGCCAAACAGTTTCTC 519
 QY 107 IleSerProAlaSerProProValGlyTyrLysGlnValGluAspAlaThrProVal 126
 DB 520 ATCTGCCCCCTTCTCCACCTGTTAGCTGGCAGCCCATCAACGATGCCACGCCAGTC 579
 QY 127 IleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyProGlyGluLysTyrGluLeu 146
 DB 580 CTCACCTATGACCTCTCTATGCTGTGGCCAACTAGACCCAGAGAGATATGAGCTC 639
 QY 147 HisAlaIleThrAspProThrProSerValValHisValCysGluSerAspGlnGlu 166
 DB 640 CATGCAGGAGCTGAGTCCACCCCAAGTGTGCTGCAGCTGGCCAGCTGCATAGAG 699
 QY 167 AsnGluGluGluGluGluMetGluArgMetLysArgProLysProLysLeuGln 186
 DB 700 GAAGAGAGGAGCCCAAG-----ACTTCCCCCAAGGCAAAATCATCCAA 744
 QY 187 ThrArgArgProGluTyrThrPro 194
 DB 745 ACTCGGCTCTGGCTGCCACCC 768

RESULT 12

US-09-782-953-20
 ; Sequence 20, Application US/09782953
 ; Patent No. US20020150953A1
 ; GENERAL INFORMATION:
 ; APPLICANT: WILLIAMS, R. SANDERS
 ; APPLICANT: ROTHERMEL, BEVERLY

; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE

; TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)
 ; FILE REFERENCE: UTSD:674PZ1
 ; CURRENT APPLICATION NUMBER: US/09/782,953
 ; CURRENT FILING DATE: 2001-02-13
 ; PRIOR APPLICATION NUMBER: 60/216,601
 ; PRIOR FILING DATE: 2000-07-07
 ; NUMBER OF SEQ ID NOS: 27
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 20
 ; LENGTH: 828
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (23)..(745)
 US-09-782-953-20

Alignment Scores:

Pred. No.: 2,386-63 Length: 828
 Score: 592.50 Matches: 120
 Percent Similarity: 71.28% Conservative: 19
 Best Local Similarity: 61.54% Mismatches: 43
 Query Match: 56.81% Indels: 13
 DB: 10 Gaps: 3

US-09-782-953-7 (1-198) x US-09-782-953-20 (1-828)

QY 1 MetAspPheArgAspPheSerTyrAsnPheSerSerLeuIleAlaCysValAlaAsnAsp 20
 DB 140 ATGGATTTAAGTATCGCT-----ACCTCATTGCTTGCAGCGTCATGAA 190
 QY 21 AspValPheSerGluSerGluThrArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40
 DB 191 GCAGTGTGTGAGGCGCAGAGCAGAGAAAGATTGAAGCACTCTTACCATCTATGAT 250
 QY 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60
 DB 251 GACCAGGTTTACTTTTTCAGCTGTTTAAAGCTTTAAGAGTCAGATAAATTCAGCAA 310
 QY 61 ProLeuSerAlaAlaAspAlaArgLeuHisLysThrGluPheLeuGlyLysGlu 80
 DB 311 CTGAAGCGCGCAGCAGAGCGCAATAGAACTCCACGAAACAGACTTCAATGGGCAGAG 370
 QY 81 MetLysLeuTyrPheAlaGlnThrLeuHisLys-----SerSerHisLeu 96
 DB 371 CTAAAGCTATATTTTGCACAGTGCAGATGTCGCGGAAGTCCGCGCAAGTCTCTATCTC 430
 QY 97 AlaProProAsnProAspLysGlnPheLeuIleSerProProAlaSerProProValGly 116
 DB 431 CTGCGCGCCCGCAGCTGTCAAGCAGTTCCTCATCTCCCTCCAGCTCTCCCGAGTGGG 490
 QY 117 TrpLysGlnValGluAspAlaThrProValIleAsnTyrAspLeuTyrAlaIleSer 136
 DB 491 TGGAAAGCAGACGAGAGATGCGATGCTGTTTATAAATATGATTTACTCTGTGCTTTCC 550
 QY 137 LysLeuGlyProGlyGluLysTyrGluLeuHisAlaIleThrAspProThrProSerVal 156
 DB 551 AAATTGGGACCAGAGAGAAATATGAATCTCCCGGGAACAGAGTCGACACCCAGCGTG 610
 QY 157 ValValHisValCysGluSerAspGlnGluAsnGluGluGluGluMetGluArg 176
 DB 611 GTGGTTCATGTCTGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 652
 QY 177 MetLysArgProLysProLysLeuIleGlnThrArgArgProGlu 191
 DB 653 ACAAAAACCCCAACAGAAAATTTGCCACAGACAGGCGCCCGAG 697

RESULT 13

US-09-782-953-23
 ; Sequence 23, Application US/09782953
 ; Patent No. US20020150953A1
 ; GENERAL INFORMATION:
 ; APPLICANT: WILLIAMS, R. SANDERS

QY 30 AlaLysPheGluSerLeuPheArgThrTyArgPheLysAspThrThrPheGlnTyrPheLys 49
 Db 410 GCCAAATTTGAGTCCCTCTTAGGAGCTATGACAGGACATCACCTTTTCAGTATTTTAAAG 351
 QY 50 SerPheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAspAlaArgLeu 69
 Db 350 AGCTTCAACGAGTCAAGATAAATTCAGCAACCCCTTCTCCGACGACAGATCCAGGCTC 291
 QY 70 ArgLeuHisLysThrGluPheLeuGlyGluMetLysLeuTyPheAlaGlnThrLeu 89
 Db 290 CAGCTGCATAAGACTGAGTTTCTGGAAAGGAATGAAGTTATATTTTTCAGGTGAGT 231
 QY 90 HisIle-----GlySerSerHisLeuAlaProAsn 100
 Db 230 TGGGTTTCATTGCTATGATGCTTCTCCCTCCCTCCCTCCCTCC 189

RESULT 15

US-09-864-761-2064/c
 ; Sequence 2064, Application US/09864761
 ; Patent No. US20020048763A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharon G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Wensheng
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 ; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
 ; FILE REFERENCE: Acomica-X-1

; CURRENT APPLICATION NUMBER: US/09/864,761

; CURRENT FILING DATE: 2001-05-23

; PRIOR APPLICATION NUMBER: US 60/180,312

; PRIOR FILING DATE: 2000-02-04

; PRIOR APPLICATION NUMBER: US 60/207,456

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: US 09/632,366

; PRIOR FILING DATE: 2000-08-03

; PRIOR APPLICATION NUMBER: GB 24263.6

; PRIOR FILING DATE: 2000-10-04

; PRIOR APPLICATION NUMBER: US 60/236,359

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: PCT/US01/00666

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00667

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00664

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00669

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00665

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00668

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00663

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00662

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00661

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00670

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: US 60/234,687

; PRIOR FILING DATE: 2000-09-21

; PRIOR APPLICATION NUMBER: US 09/608,408

; PRIOR FILING DATE: 2000-06-30

; PRIOR APPLICATION NUMBER: US 09/774,203

; PRIOR FILING DATE: 2001-01-29

NUMBER OF SEQ ID NOS: 49117

SOFTWARE: Annomax Sequence Listing Engine vers. 1.1

SEQ ID NO 2064

LENGTH: 446

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: MAP TO AP000122.1
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7
 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.1
 OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 4.1
 OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.7
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7
 OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.5
 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.3
 OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.2
 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2
 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.4
 US-09-864-761-2064

Alignment Scores:

Pred. No.: 3,1e-26 Length: 446
 Score: 285.50 Matches: 58
 Percent Similarity: 82.43% Conservatives: 3
 Best Local Similarity: 78.38% Mismatches: 10
 Query Match: 27.37% Indels: 3
 DB: 10 Gaps: 1

US-09-782-953-7 (1-198) x US-09-864-761-2064 (1-446)

QY 30 AlaLysPheGluSerLeuPheArgThrTyArgPheLysAspThrThrPheGlnTyrPheLys 49
 Db 444 GCCAAATTTGAGTCCCTCTTAGGAGCTATGACAGGACATCACCTTTTCAGTATTTTAAAG 385
 QY 50 SerPheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAspAlaArgLeu 69
 Db 384 AGCTTCAACGAGTCAAGATAAATTCAGCAACCCCTTCTCCGACGACAGATCCAGGCTC 325
 QY 70 ArgLeuHisLysThrGluPheLeuGlyGluMetLysLeuTyPheAlaGlnThrLeu 89
 Db 324 CAGCTGCATAAGACTGAGTTTCTGGAAAGGAATGAAGTTATATTTTTCAGGTGAGT 265
 QY 90 HisIle-----GlySerSerHisLeuAlaProAsn 100
 Db 264 TGGGTTTCATTGCTATGATGCTTCTCCCTCCCTCCCTCCCTCC 223

Search completed: December 15, 2002, 00:06:25

Job time : 38.0556 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2002, 15:09:05 ; Search time 1042.77 Seconds
(without alignments)
14442.598 Million cell updates/sec

Title: US-09-782-953-1

Perfect score: 599

Sequence: 1 gagggcaaggaaactcca.....gaggagggaaggaagagat 599

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 24791104 seqs, 12571243825 residues

Total number of hits satisfying chosen parameters: 49582208

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents NA Main:*

- 1: /cgn2_6/ptodata/1/pna/pctus COMB seq:*
- 2: /cgn2_6/ptodata/1/pna/US06 COMB seq:*
- 3: /cgn2_6/ptodata/1/pna/US07 COMB seq:*
- 4: /cgn2_6/ptodata/1/pna/US08 COMB seq:*
- 5: /cgn2_6/ptodata/1/pna/US081 COMB seq:*
- 6: /cgn2_6/ptodata/1/pna/US082 COMB seq:*
- 7: /cgn2_6/ptodata/1/pna/US083 COMB seq:*
- 8: /cgn2_6/ptodata/1/pna/US084 COMB seq:*
- 9: /cgn2_6/ptodata/1/pna/US085 COMB seq:*
- 10: /cgn2_6/ptodata/1/pna/US086 COMB seq:*
- 11: /cgn2_6/ptodata/1/pna/US087 COMB seq:*
- 12: /cgn2_6/ptodata/1/pna/US088 COMB seq:*
- 13: /cgn2_6/ptodata/1/pna/US089 COMB seq:*
- 14: /cgn2_6/ptodata/1/pna/US090 COMB seq:*
- 15: /cgn2_6/ptodata/1/pna/US091 COMB seq:*
- 16: /cgn2_6/ptodata/1/pna/US092 COMB seq:*
- 17: /cgn2_6/ptodata/1/pna/US093 COMB seq:*
- 18: /cgn2_6/ptodata/1/pna/US094 COMB seq:*
- 19: /cgn2_6/ptodata/1/pna/US095 COMB seq:*
- 20: /cgn2_6/ptodata/1/pna/US095A COMB seq:*
- 21: /cgn2_6/ptodata/1/pna/US095B COMB seq:*
- 22: /cgn2_6/ptodata/1/pna/US095C COMB seq:*
- 23: /cgn2_6/ptodata/1/pna/US096 COMB seq:*
- 24: /cgn2_6/ptodata/1/pna/US096B COMB seq:*
- 25: /cgn2_6/ptodata/1/pna/US096C COMB seq:*
- 26: /cgn2_6/ptodata/1/pna/US096D COMB seq:*
- 27: /cgn2_6/ptodata/1/pna/US096E COMB seq:*
- 28: /cgn2_6/ptodata/1/pna/US097A COMB seq:*
- 29: /cgn2_6/ptodata/1/pna/US097B COMB seq:*
- 30: /cgn2_6/ptodata/1/pna/US097C COMB seq:*
- 31: /cgn2_6/ptodata/1/pna/US098A COMB seq:*
- 32: /cgn2_6/ptodata/1/pna/US098B COMB seq:*
- 33: /cgn2_6/ptodata/1/pna/US098C COMB seq:*
- 34: /cgn2_6/ptodata/1/pna/US099A COMB seq:*
- 35: /cgn2_6/ptodata/1/pna/US099B COMB seq:*
- 36: /cgn2_6/ptodata/1/pna/US099C COMB seq:*
- 37: /cgn2_6/ptodata/1/pna/US099D COMB seq:*
- 38: /cgn2_6/ptodata/1/pna/US100A COMB seq:*
- 39: /cgn2_6/ptodata/1/pna/US100B COMB seq:*
- 40: /cgn2_6/ptodata/1/pna/US101A COMB seq:*
- 41: /cgn2_6/ptodata/1/pna/US101B COMB seq:*
- 42: /cgn2_6/ptodata/1/pna/US102A COMB seq:*
- 43: /cgn2_6/ptodata/1/pna/US102B COMB seq:*

- 44: /cgn2_6/ptodata/1/pna/US6000 COMB seq:*
- 45: /cgn2_6/ptodata/1/pna/US6001 COMB seq:*
- 46: /cgn2_6/ptodata/1/pna/US6002 COMB seq:*
- 47: /cgn2_6/ptodata/1/pna/US6003 COMB seq:*
- 48: /cgn2_6/ptodata/1/pna/US6004 COMB seq:*
- 49: /cgn2_6/ptodata/1/pna/US6005 COMB seq:*
- 50: /cgn2_6/ptodata/1/pna/US6006 COMB seq:*
- 51: /cgn2_6/ptodata/1/pna/US6007 COMB seq:*
- 52: /cgn2_6/ptodata/1/pna/US6008 COMB seq:*
- 53: /cgn2_6/ptodata/1/pna/US6009 COMB seq:*
- 54: /cgn2_6/ptodata/1/pna/US6010 COMB seq:*
- 55: /cgn2_6/ptodata/1/pna/US6011 COMB seq:*
- 56: /cgn2_6/ptodata/1/pna/US6012 COMB seq:*
- 57: /cgn2_6/ptodata/1/pna/US6013 COMB seq:*
- 58: /cgn2_6/ptodata/1/pna/US6014 COMB seq:*
- 59: /cgn2_6/ptodata/1/pna/US6015 COMB seq:*
- 60: /cgn2_6/ptodata/1/pna/US6016 COMB seq:*
- 61: /cgn2_6/ptodata/1/pna/US6017 COMB seq:*
- 62: /cgn2_6/ptodata/1/pna/US6018 COMB seq:*
- 63: /cgn2_6/ptodata/1/pna/US6019 COMB seq:*
- 64: /cgn2_6/ptodata/1/pna/US6020 COMB seq:*
- 65: /cgn2_6/ptodata/1/pna/US6021 COMB seq:*
- 66: /cgn2_6/ptodata/1/pna/US6022 COMB seq:*
- 67: /cgn2_6/ptodata/1/pna/US6023 COMB seq:*
- 68: /cgn2_6/ptodata/1/pna/US6024 COMB seq:*
- 69: /cgn2_6/ptodata/1/pna/US6025 COMB seq:*
- 70: /cgn2_6/ptodata/1/pna/US6026 COMB seq:*
- 71: /cgn2_6/ptodata/1/pna/US6027 COMB seq:*
- 72: /cgn2_6/ptodata/1/pna/US6028 COMB seq:*
- 73: /cgn2_6/ptodata/1/pna/US6029 COMB seq:*
- 74: /cgn2_6/ptodata/1/pna/US6030 COMB seq:*
- 75: /cgn2_6/ptodata/1/pna/US6031 COMB seq:*
- 76: /cgn2_6/ptodata/1/pna/US6032 COMB seq:*
- 77: /cgn2_6/ptodata/1/pna/US6033 COMB seq:*
- 78: /cgn2_6/ptodata/1/pna/US6034 COMB seq:*
- 79: /cgn2_6/ptodata/1/pna/US6035 COMB seq:*
- 80: /cgn2_6/ptodata/1/pna/US6036 COMB seq:*
- 81: /cgn2_6/ptodata/1/pna/US6037 COMB seq:*
- 82: /cgn2_6/ptodata/1/pna/US6038 COMB seq:*
- 83: /cgn2_6/ptodata/1/pna/US6039 COMB seq:*
- 84: /cgn2_6/ptodata/1/pna/US6040 COMB seq:*
- 85: /cgn2_6/ptodata/1/pna/US6041 COMB seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	599	100.0	599	30	US-09-782-953-1
2	554	92.5	592	28	US-09-704-424-23327
3	504.4	84.2	597	30	US-09-782-953-5
4	500	83.5	538	28	US-09-704-424-23248
5	427.8	71.4	597	80	US-60-360-207-8520
6	427.4	71.4	613	18	US-09-401-645-4194
7	426.2	71.2	597	22	US-09-575-580B-2
8	426.2	71.2	597	30	US-09-782-953-2
9	413.8	69.1	1880	18	US-09-432-241A-4522
10	413.8	69.1	2331	30	US-09-782-953-11
11	413.8	69.1	2340	42	US-10-247-671-100
12	413.8	69.1	2340	76	US-60-323-784-100
13	413.8	69.1	2346	1	PCT-US02-02176-1165
14	413.8	69.1	2346	1	PCT-US02-23766-4321
15	413.8	69.1	2355	38	US-10-044-090-255
16	413.8	69.1	2355	39	US-10-084-817-36
17	413.8	69.1	2355	40	US-10-116-802-289
18	413.8	69.1	2355	70	US-60-260-483-255
19	413.8	69.1	2355	72	US-60-281-593-289
20	412.2	68.8	580	25	US-09-644-863-7097
21	410	68.4	456	18	US-09-436-762A-33053

```

22 405.8 67.7 594 25 US-09-652-355-3836 Sequence 1836, Ap
23 405 67.6 440 18 US-09-436-762A-28336 Sequence 28336, A
24 404.8 67.6 2348 63 US-09-195-106-60 Sequence 60, Ap1
25 400.8 67.2 581 28 US-09-710-280-1577 Sequence 1577, Ap1
26 400.8 66.9 567 20 US-09-539-802A-533 Sequence 533, App
27 399.6 66.7 619 18 US-09-401-645-9836 Sequence 9836, Ap
28 394.4 65.8 593 29 US-09-726-787-1970 Sequence 1970, Ap
29 384 64.1 578 28 US-09-710-280-1000 Sequence 1000, Ap
30 377 62.9 401 18 US-09-436-762A-16620 Sequence 16620, A
31 374.4 62.5 972 17 US-09-359-067-31161 Sequence 31161, A
32 373 62.3 397 18 US-09-436-762A-18907 Sequence 18907, A
33 370.6 61.9 518 25 US-09-652-128-4690 Sequence 4690, Ap
34 369.4 61.7 586 29 US-09-726-787-1963 Sequence 1963, Ap
35 369 61.6 393 18 US-09-436-762A-31396 Sequence 31396, A
36 368 61.4 541 20 US-09-539-802A-842 Sequence 842, App
37 360.8 60.2 583 25 US-09-644-869-7149 Sequence 7149, Ap
38 360 60.1 566 25 US-09-644-869-6053 Sequence 6053, Ap
39 359.8 60.1 540 29 US-09-726-787-1098 Sequence 1098, Ap
40 357.8 59.7 2358 14 PCT-US00-05918-347 Sequence 347, App
41 357.8 59.7 2358 34 US-09-925-302-347 Sequence 347, App
42 357.4 59.7 524 25 US-09-652-123-3252 Sequence 3252, Ap
43 357.4 59.7 562 20 US-09-539-802A-2692 Sequence 2692, Ap
44 356.4 58.5 552 28 US-09-710-280-1359 Sequence 1359, Ap
45 350.6 58.5 563 28 US-09-716-475-4375 Sequence 4375, Ap

```

ALIGNMENTS

```

RESULT 1
US-09-782-953-1
; Sequence 1, Application US/09782953
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: US/09-782-953
; CURRENT APPLICATION NUMBER: US/09-782-953
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216, 601
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 599
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-782-953-1

```

```

Query Match 100.0%; Score 599; DB 30; Length 599;
Best Local Similarity 100.0%; Pred. No. 5,5e-172;
Matches 599; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GAGGTGCAAGAACTCCAGCTTGGGCTTGAAGAGAGCCAGTCTGTTAGCTTACCGT 60
DB 1 GAGGTGCAAGAACTCCAGCTTGGGCTTGAAGAGAGCCAGTCTGTTAGCTTACCGT 60
QY 61 CTGCCCCGTAAGAAAGAGAAATATTTAGGACTTATAGCTTACCAATTTAGCTCCCGAT 120
DB 61 CTGCCCCGTAAGAAAGAGAAATATTTAGGACTTATAGCTTACCAATTTAGCTCCCGAT 120
QY 121 TGGTGTGTGGCAAAAGATGATCTTCAAGCAAGAGTGAAGCCAGGCAATTTGATC 180
DB 121 TGGTGTGTGGCAAAAGATGATCTTCAAGCAAGAGTGAAGCCAGGCAATTTGATC 180
QY 181 CTTCTTGAAGCATATATGCAAGAGACACCTTCAATATTTTAAGAGCTTCAAACTGT 240
DB 181 CTTCTTGAAGCATATATGCAAGAGACACCTTCAATATTTTAAGAGCTTCAAACTGT 240
QY 241 CCGGATAAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
DB 241 CCGGATAAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300

```

```

QY 301 CGAGTCTCGGGAGAGAAATGAATGTATTTTGTCTCAAGCTTTACATAGAACTTC 360
DB 301 CGAGTCTCGGGAGAGAAATGAATGTATTTTGTCTCAAGCTTTACATAGAACTTC 360
QY 361 ACACCTGGCTCCCGCATTCGCAACAAACAGTTCTCTCTCCCTCGGCTCCCTCC 420
DB 361 ACACCTGGCTCCCGCATTCGCAACAAACAGTTCTCTCTCCCTCGGCTCCCTCC 420
QY 421 GTTGGCTGAAACAGTAGAAGATGCCACCCGCTCAATTAATAGATCTTTATATGCC 480
DB 421 GTTGGCTGAAACAGTAGAAGATGCCACCCGCTCAATTAATAGATCTTTATATGCC 480
QY 481 ATTCCTCAAGCTGGGCGCCAGAGAGAGATGAACTGATGAGAGACACCACTCC 540
DB 481 ATTCCTCAAGCTGGGCGCCAGAGAGAGATGAACTGATGAGAGACACCACTCC 540
QY 541 AGTGTGTGTCTCCACGTGTGTGAGAGTGAACCAAGAAATGAGAGAGAGAGAT 599
DB 541 AGTGTGTGTCTCCACGTGTGTGAGAGTGAACCAAGAAATGAGAGAGAGAGAT 599

```

```

RESULT 2
US-09-704-424-23327
; Sequence 23327, Application US/09704424
; GENERAL INFORMATION:
; APPLICANT: Watson, James
; TITLE OF INVENTION: Polynucleotides, Polypeptides Expressed
; FILE REFERENCE: 11000.1045U
; CURRENT APPLICATION NUMBER: US/09-704,424
; CURRENT FILING DATE: 2000-11-01
; NUMBER OF SEQ ID NOS: 36881
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23327
; LENGTH: 592
; TYPE: DNA
; ORGANISM: Mouse
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(592)
; OTHER INFORMATION: n = A,T,C or G
US-09-704-424-23327

```

```

Query Match 92.5%; Score 554; DB 28; Length 592;
Best Local Similarity 99.3%; Pred. No. 3.3e-158;
Matches 587; Conservative 0; Mismatches 1; Indels 3; Gaps 3;

```

```

QY 11 GGAACCTCCAGCTTGGGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 70
DB 1 GGAACCTCCAGCTTGGGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 70
QY 71 AAAAAGCAAGATG-ATTTAGGAGATTTAGCTTACAAATTTAGCTTCTGATGCTTGT 129
DB 71 AAAAAGCAAGATGATTTAGGAGATTTAGCTTACAAATTTAGCTTCTGATGCTTGT 129
QY 130 GGCACCAAGATGATGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 189
DB 130 GGCACCAAGATGATGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 189
QY 190 AACATATGCAAGAGACACCTTCAATATTTTAAGAGCTTCAAGAGAGAGAGAG 249
DB 190 AACATATGCAAGAGACACCTTCAATATTTTAAGAGCTTCAAGAGAGAGAGAG 249
QY 250 CTTGAGCAACCCCTTATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 309
DB 250 CTTGAGCAACCCCTTATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 309
QY 310 GGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 369
DB 310 GGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 369
QY 370 TCCG-CCAAATCCGCAAAACAGTTCTCTATCTCCCTCGGCTCTCTCTCGGCTG 428
DB 370 TCCG-CCAAATCCGCAAAACAGTTCTCTATCTCCCTCGGCTCTCTCTCGGCTG 428

```



```

; Sequence 8520, Application US/60360207
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig
; TITLE OF INVENTION: HUMAN GENOME DISCOVERY SYSTEM AND USES THEREOF
; FILE REFERENCE: CL001321
; CURRENT APPLICATION NUMBER: US/60/360,207
; CURRENT FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 47235
; SEQ ID NO 8520
; LENGTH: 597
; TYPE: DNA
; ORGANISM: HUMAN
US-60-360-207-8520

```

Query Match 71.4%; Score 427.8; DB 80; Length 597;

Best Local Similarity 93.3%; Pred. No. 1,4e-119; Mismatches 32; Indels 1; Gaps 1;

```

Matches 458; Conservative 0; Mismatches 32; Indels 1; Gaps 1;

QY 110 AGCTCCGATGCTGTGTGGCAACGATGATGCTTCAGCAAGAGAGACCAAGGCC 169
DB 31 AGGCCACCATGCTGCTGCCACCTGACCGCGCTGTGTGAGAGCGCTGTGCGGCC 90
QY 170 AATTGAAATCCCTCTTCAGAACATATGACCAAGACCACTTCAGATTTTAAAGC 229
DB 91 AATTGAAATCCCTCTTCAGAACATATGACCAAGACCACTTCAGATTTTAAAGC 150
QY 230 TTCAAAAGTGTGGGATTAACCTTCACCAACCCCTTATCTGACGCGATGCGAGCGG 289
DB 151 TTCAAAAGTGTGGGATTAACCTTCACCAACCCCTTATCTGACGCGATGCGAGCGG 210
QY 290 CTCACAAAGCCAGTTCCTGCGGAAAGAAATGATGATTTTCTGACATTTTAC 349
DB 211 CTCACAAAGCCAGTTCCTGCGGAAAGAAATGATGATTTTCTGACATTTTAC 270
QY 350 ATAGGAAGTTCAACCTGCTCCG-CCAATCCGCAAAAGTTCTCATTCCTCCG 408
DB 271 ATAGGAAGTTCAACCTGCTCCG-CCAATCCGCAAAAGTTCTCATTCCTCCG 330
QY 409 GCCTCTCCCTCCGCTGGGAAAGAAAGATGATGATGATGATGATGATGATGAT 468
DB 331 GCCTCTCCCTCCGCTGGGAAAGAAAGATGATGATGATGATGATGATGATGAT 390
QY 469 CTTTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 528
DB 391 CTTTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 450
QY 529 GACACCACTCCCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 588
DB 451 GACACCACTCCCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 510
QY 589 GAGGAGAGAT 599
DB 511 GAGGAGAGAT 521

```

RESULT 6

US-09-401-645-4194

```

; Sequence 4194, Application US/09401645
; GENERAL INFORMATION:
; APPLICANT: Geating, David P.
; APPLICANT: Diselano, Peter
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
; TITLE OF INVENTION: RAT SKELETAL MUSCLE LIBRARY
; FILE REFERENCE: MN98-43DA
; CURRENT APPLICATION NUMBER: US/09/401,645
; CURRENT FILING DATE: 1999-09-22
; EARLIER APPLICATION NUMBER: 60/101,346
; EARLIER FILING DATE: 1998-09-22
; NUMBER OF SEQ ID NOS: 4253
; SOFTWARE: ParseSeq for Windows Version 3.0
; SEQ ID NO 4194
; LENGTH: 613

```

```

; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(613)
; OTHER INFORMATION: n = A,T,C or G
US-09-401-645-4194

```

Query Match 71.4%; Score 427.4; DB 18; Length 613;
Best Local Similarity 85.1%; Pred. No. 1,9e-119;
Matches 504; Conservative 0; Mismatches 84; Indels 4; Gaps 3;

```

QY 9 AAGAACCTCCAGCTTGGCTTGA-AGAGAGGAGTCTTCTTGAAGCTTGGCC 66
DB 1 AAGAACCTCCAGCTTGGCTTGA-AGAGAGGAGTCTTCTTGAAGCTTGGCC 60
QY 67 CGTGAAGAGAGAGT-ATTTAGGACTTTAGCTTCAATTTTAACTCCCTGATGCTT 125
DB 61 GTTGAAGAGAGAGT-ATTTAGGACTTTAGCTTCAATTTTAACTCCCTGATGCTT 120
QY 126 GTGTGGCAACGATGATGCTTCAGCGAAAGTGAAGACAGGCGCAATTTGAATCCCTCT 185
DB 121 GTGTGGCAACGATGATGCTTCAGCGAAAGTGAAGACAGGCGCAATTTGAATCCCTCT 180
QY 186 TCAGAACATATGACAGGACCACTTCAGATTTTAAAGCTTCAAGCTGTCGGA 245
DB 181 TCAGAACATATGACAGGACCACTTCAGATTTTAAAGCTTCAAGCTGTCGGA 239
QY 246 TAACTTCAGCAACCCCTTATCTGACGCGATGCGAGCTGCGCTGACAGACGAGT 305
DB 240 TAACTTCAGCAACCCCTTATCTGACGCGATGCGAGCTGCGCTGACAGACGAGT 299
QY 306 TCCTGGGAGAGAAATGAAGTTGATTTTGTCTCAAGCTTTACATGAGAAAGTTCAACC 365
DB 300 TCCTGGGAGAGAAATGAAGTTGATTTTGTCTCAAGCTTTACATGAGAAAGTTCAACC 359
QY 366 TGGCTCCGCAATCCCGCAAAAGTTCTCATCTCCGCTCTCTCCGCTTGG 425
DB 360 TGGCTCCGCAATCCCGCAAAAGTTCTCATCTCCGCTCTCTCCGCTTGG 419
QY 426 CTGGAACAGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 485
DB 420 CTGGAACAGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 479
QY 486 CAAGCTGGGCGGAGAGAAATGAACTGATGACAGGACAGACCACTCCAGGT 545
DB 480 CAAGCTGGGCGGAGAGAAATGAACTGATGACAGGACAGACCACTCCAGGT 539
QY 546 GGTGGTCCAGTGTGTGAGATGACCAAGAGATGAGAGAGAGAGAGAG 597
DB 540 GGTGGTCCAGTGTGTGAGATGACCAAGAGATGAGAGAGAGAGAGAGAG 591

```

RESULT 7

US-09-575-580B-2

```

; Sequence 2, Application US/09575580B
; GENERAL INFORMATION:
; APPLICANT: McKee, P.
; APPLICANT: Kayako, K.
; APPLICANT: Ryeom, S.
; TITLE OF INVENTION: CALCIPRESSINS: ENDOGENOUS INHIBITORS OF CALCINEURIN
; TITLE OF INVENTION: USES AND REAGENTS RELATED THERETO
; FILE REFERENCE: HMV-048.01
; CURRENT APPLICATION NUMBER: US/09/575,580B
; CURRENT FILING DATE: 2000-05-22
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: Patencin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 597
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-575-580B-2

```

QY	170	AAATTTGAATCCCTCTTCAGAACATATGACAAGGACACCACTTCCAGTATTTTAAAGAGC	229
Db	91	AAATTTGAATCCCTCTTCAGAACATATGACAAGGACACCACTTCCAGTATTTTAAAGAGC	150
QY	230	TTCAAAACGTGTCGGGATAAATCTCAGCAACCCCTTATCTCGAGCCGATGCCAGGCTGCGG	289
Db	151	TTCAAAACGTGTCGGGATAAATCTCAGCAACCCCTTATCTCGAGCCGATGCCAGGCTGCGG	210
QY	290	CTGCAACAAGCCGAGTTCCTGGGGAAGAAATCAAGTTGTATTTTGTCTCAGACTTTTACAC	349
Db	211	CTGCAACAAGCCGAGTTCCTGGGGAAGAAATCAAGTTGTATTTTGTCTCAGACTTTTACAC	270
QY	350	ATAGGAAGTTCACACCTGGCTCCG-CCAATCCGACAAACAGTTCTCATCTCCCTCCG	408
Db	271	ATAGGAAGTTCACACCTGGCTCCG-CCAATCCGACAAACAGTTCTCATCTCCCTCCG	330
QY	409	GCCTCTCTCCCTGGCTGGAAACAAGTAGAAGATGCCACCCCTGTCAATAAATTACGAT	468
Db	331	GCCTCTCTCCCTGGCTGGAAACAAGTAGAAGATGCCACCCCTGTCAATAAATTACGAT	390
QY	469	CTTTTATATGCCATCTCCAAGCTGGGGCCAGGAGAGAAGTATGAATCGATCGAGCGACA	528
Db	391	CTTTTATATGCCATCTCCAAGCTGGGGCCAGGAGAGAAGTATGAATCGATCGAGCGACA	450
QY	529	GACACCACTCCAGTGTGTGTGTCACAGTGTGTGAGAGTGACCAAGAGAATCAGAGGAA	588
Db	451	GACCCCACTCCAGTGTGTGTGTCACAGTGTGTGAGAGTGACCAAGAGAATCAGAGGAA	510
QY	589	GAGGAAGAGAT	599
Db	511	GAGGAAGAGAT	521
RESULT 9			
US-09-432-241A-4522			
; Sequence 4522, Application US/09432241A			
; GENERAL INFORMATION:			
; APPLICANT: Gearing, David P.			
; APPLICANT: Holtzman, Douglas A.			
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A			
; TITLE OF INVENTION: HUMAN EPIDERMAL KERATINOCYTE LIBRARY			
; FILE REFERENCE: 1600.1004001			
; CURRENT APPLICATION NUMBER: US/09/432,241A			
; CURRENT FILING DATE: 1999-10-28			
; PRIOR APPLICATION NUMBER: 60/106,445			
; PRIOR FILING DATE: 1998-10-30			
; PRIOR APPLICATION NUMBER: 60/107,227			
; PRIOR FILING DATE: 1998-11-05			
; PRIOR APPLICATION NUMBER: 60/127,182			
; PRIOR FILING DATE: 1999-03-30			
; PRIOR APPLICATION NUMBER: 60/132,099			
; PRIOR FILING DATE: 1999-04-30			
; NUMBER OF SEQ ID NOS: 5041			
; SOFTWARE: FastSeq for Windows Version 3.0			
; SEQ ID NO 4522			
; LENGTH: 1880			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-09-432-241A-4522			
Query Match			
Best Local Similarity 69.1%; Score 413.8; DB 18; Length 1880;			
Matches 474; Conservative 0; Mismatches 82; Indels 1; Gaps			
QY	39	GAGCGAGTCGTTTCGTTAGCGCTCTCCCGTGAAGAGCAAGATGATTTTAGGCACTTTA	98
Db	89	GACTCGTGGGTCTGTAGCGCTTTCTACTGTAAGAAAGCAAGATGCAATTTTAGAAACTTTA	148
QY	99	GCTACAAATTTTAGCTCCCTGATTTGCTGTGTGGCAACAGATGATGCTTCAGCGAAATG	158
Db	149	ACTACAGTTTTAGCTCCCTGATTCCTGTGTGGCAACAGATGATGCTTCAGCGAAATG	208
QY	159	AGACACGGGCCAAATTTGAATCCCTCTTCAGAACATATGACAAGGACACCACTTCCAGT	218

```

Db 209 AACCCAGGCGCAATTGAGTCCCTCTTTAGAGCATGACAGACATCACCTTTAGT 268
Qy 219 ATTTTAAGAGCTTCAAAAGTGTCCGSAATTAACCTTCAGCAACCCCTTATCTGCAGCCGATG 278
Db 269 ATTTTAAGAGCTTCAAAAGTGTCCGSAATTAACCTTCAGCAACCCCTTATCTGCAGCCGATG 328
Qy 279 CCAGGCTGCGGCTGCACAAAGCCGAGTTCTCTGGGAAAGAAATGAATGTATTTTCTC 338
Db 329 CCAGGCTGCGGCTGCACAAAGCCGAGTTCTCTGGGAAAGAAATGAATGTATTTTCTC 388
Qy 339 AGACTTTACATAGGAAGTTCAACCTGCTCCGCC-ATCCGACAAACAGTTCTCTCA 397
Db 389 AGACTTTACATAGGAAGTTCAACCTGCTCCGCC-ATCCGACAAACAGTTCTCTCA 448
Qy 398 TCTCCCTCCGCGCTCTCTCCGCTGGCTGGAAACAAATGAAAGTCCACCCCTCA 457
Db 449 TCTCCCTCCGCGCTCTCTCCGCTGGGATGGAAACAAATGAAAGTCCACCCCTCA 508
Qy 458 TAAATTACGATCTTTATATGTCATCTCCAGCTGGGCGCAGAGAGAAATGAACTGC 517
Db 509 TAAATTACGATCTTTATATGTCATCTCCAGCTGGGCGCAGAGAGAAATGAAATTTGC 568
Qy 518 ATGACGCGACAGACACCACTCCAGTGTGTGTGTCCAGTGTGTGAGAGTGCACCAAGAGA 577
Db 569 ACGCAGCGACTGACACCACTCCAGCTGTGTGTGTCCAGTGTGTGAGAGTGCATCAAGAGA 628
Qy 578 ATGAGAGAGAGAGGAA 594
Db 629 AGAGAGAGAGAGGAA 645

```

```

RESULT 10
US-09-782-953-11
; Sequence 11, Application US/09782953
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: US/09-782-953
; CURRENT FILING DATE: 2001-02-13
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,601
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 11
; LENGTH: 2331
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (144)..(734)
US-09-782-953-11

```

```

Query Match 69.1%; Score 413.8; DB 30; Length 2331;
Best Local Similarity 85.1%; Pred. No. 4.8e-115;
Matches 474; Conservative 0; Mismatches 82; Indels 1; Gaps 1;
Qy 39 GAGCGAGTCTGTTGTTAAGCTGTGCGCGTGAAGAAAGCAAGATGATTTTAGGACCTTTA 98
Db 103 GACTGCGTGGGTCTGTAGCCCTTCACTGAAGAAACAGATGCAATTTTAGAACTTTA 162
Qy 99 GCTACAAATTTTAGTCCCTGATGCTGTGTGCGAAACGATGATGTTTACGGAAGTG 158
Db 163 ACTACAGTTTAACTCTCTGATGCTGTGTGTGCGAAACGATGATGTTTACGGAAGTG 222
Qy 159 AGACGAGGCGCAAAATTTGAATCTCTTTCAGAACATATGACAAAGACCAACCTTCAGT 218
Db 223 AAACGAGGCGCAAAATTTGAATCTCTTTCAGAACATATGACAAAGACCAACCTTCAGT 282
Qy 219 ATTTTAAGAGCTTCAAAAGTGTCCGSAATTAACCTTCAGCAACCCCTTATCTGCAGCCGATG 278

```

```

Db 283 ATTTTAAGAGCTTCAAAAGTGTCCGSAATTAACCTTCAGCAACCCCTTATCTGCAGCCGATG 342
Qy 279 CCAGGCTGCGGCTGCACAAAGCCGAGTTCTCTGGGAAAGAAATGAATGTATTTTCTC 338
Db 343 CCAGGCTGCGGCTGCACAAAGCCGAGTTCTCTGGGAAAGAAATGAATGTATTTTCTC 402
Qy 339 AGACTTTACATAGGAAGTTCAACCTGCTCCGCC-ATCCGACAAACAGTTCTCTCA 397
Db 403 AGACTTTACATAGGAAGTTCAACCTGCTCCGCC-ATCCGACAAACAGTTCTCTCA 462
Qy 398 TCTCCCTCCGCGCTCTCTCCGCTGGCTGGAAACAAATGAAAGTCCACCCCTCA 457
Db 463 TCTCCCTCCGCGCTCTCTCCGCTGGGATGGAAACAAATGAAAGTCCACCCCTCA 522
Qy 458 TAAATTACGATCTTTATATGTCATCTCCAGCTGGGCGCAGAGAGAAATGAACTGC 517
Db 523 TAAATTACGATCTTTATATGTCATCTCCAGCTGGGCGCAGAGAGAAATGAAATTTGC 582
Qy 518 ATGACGCGACAGACACCACTCCAGTGTGTGTGTCCAGTGTGTGAGAGTGCACCAAGAGA 577
Db 583 ACGCAGCGACTGACACCACTCCAGCTGTGTGTGTCCAGTGTGTGAGAGTGCATCAAGAGA 642
Qy 578 ATGAGAGAGAGAGGAA 594
Db 643 AGAGAGAGAGAGGAA 659

```

```

RESULT 11
US-10-247-671-100
; Sequence 100, Application US/10247671
; GENERAL INFORMATION:
; APPLICANT: Mikita, Thomas
; APPLICANT: Shiffman, Dov
; APPLICANT: Porter, Gordon, J.
; APPLICANT: Kasey, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS
; FILE REFERENCE: PA-0050 US
; CURRENT FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/323,784
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PERL Program
; SEQ ID NO 100
; LENGTH: 2340
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incycle ID No: 5511889CBI
US-10-247-671-100

```

```

Query Match 69.1%; Score 413.8; DB 42; Length 2340;
Best Local Similarity 85.1%; Pred. No. 4.8e-115;
Matches 474; Conservative 0; Mismatches 82; Indels 1; Gaps 1;
Qy 39 GAGCGAGTCTGTTGTTAAGCTGTGCGCGTGAAGAAAGCAAGATGATTTTAGGACCTTTA 98
Db 110 GACTGCGTGGGTCTGTAGCCCTTCACTGAAGAAACAGATGCAATTTTAGAACTTTA 169
Qy 99 GCTACAAATTTTAGTCCCTGATGCTGTGTGCGAAACGATGATGTTTACGGAAGTG 158
Db 170 ACTACAGTTTAACTCTCTGATGCTGTGTGTGCGAAACGATGATGTTTACGGAAGTG 229
Qy 159 AGACGAGGCGCAAAATTTGAATCTCTTTCAGAACATATGACAAAGACCAACCTTCAGT 218
Db 230 AAACGAGGCGCAAAATTTGAATCTCTTTCAGAACATATGACAAAGACCAACCTTCAGT 289
Qy 219 ATTTTAAGAGCTTCAAAAGTGTCCGSAATTAACCTTCAGCAACCCCTTATCTGCAGCCGATG 278
Db 290 ATTTTAAGAGCTTCAAAAGTGTCCGSAATTAACCTTCAGCAACCCCTTATCTGCAGCCGATG 349

```

QY 279: CCAGGCTCGGCTGCACAGACCGAGTTCTCGGGGAAGAAATGAAGTTGTATTTTGCTC 338
 Db |||||
 QY 350 CCAGGCTCCAGCTGCATAGACTGAGTTTCTGGGAAGAAATGAAGTTATATTTTGCTC 409
 Db |||||
 QY 339 AGACTTTTACACATAGGAAGTTCCACACCTGGCTCCGCC-AAATCCCGACAAACAGATTCCTCA 397
 Db |||||
 QY 410 AGACCTTTACACATAGGAAGTTCCACACCTGGCTCCGCCAAATCCAGACAAAGCAGTTTCTGA 469
 Db |||||
 QY 398 TCTCCCTCCCGCTCTCTCCCTGGTGGTGGAAACAAAGTAGAAGATGCCACCCCGTCA 457
 Db |||||
 QY 470 TCTCCCTCCCGCTCTCCCGAGTGGATGGAAACAAGTGGAGATCGGACCCAGTCA 529
 Db |||||
 QY 458 TAAATTTACGATCTTTTATATGCTCTCCAGCTGGGCGGACGAGAGAAAGTATGAACCTGC 517
 Db |||||
 QY 530 TAAACTATGATCTCTTATATGCTCTCCAGCTGGGCGGACGAGAGAAAGTATGAATTCG 589
 Db |||||
 QY 518 ATGCAGCGACAGACACCACTCCAGTGTGGTGGTCCAGCTGTGTGAGAGTGACCAAGAGA 577
 Db |||||
 QY 590 ACGCAGCGACTGACACCACTCCAGCGTGTGGTCCATGTATGTGAGAGTGTCAAGAGA 649
 Db |||||
 QY 578 ATGAGGAGGAGAGGAA 594
 Db |||||
 QY 650 AGGAGGAAGAGAGGAA 666
 Db |||||

RESULT 12

US-60-323-784-100

; Sequence 100, Application US/60323784

; GENERAL INFORMATION:

; APPLICANT: Mikita, Thomas

; APPLICANT: Shiffman, Dov

; APPLICANT: Porter, Gordon, J.

; APPLICANT: Kaer, Matthew R.

; TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS

; FILE REFERENCE: PA-0050 P

; CURRENT APPLICATION NUMBER: US/60/323,784

; CURRENT FILING DATE: 2001-09-19

; NUMBER OF SEQ ID NOS: 186

; SOFTWARE: PERL Program

; SEQ ID NO 100

; LENGTH: 2340

; TYPE: DNA

; ORGANISM: Homo sapiens

; NAME/KEY: misc feature

; OTHER INFORMATION: Incyte ID No: 5511899CBI

US-60-323-784-100

Query Match 69.1%; Score 413.8; DB 76; Length 2340;
 Best Local Similarity 85.1%; Pred. No. 4.8e-115;
 Matches 474; Conservative 0; Mismatches 82; Indels 1; Gaps 1;

QY 39 GAGCGAGTCGTTGTTAAGCGTCTGCCCGTGAAAGACGAGAAATGATTTTAGGACTTTA 98
 Db |||||
 QY 110 GACTGCTGGGTCTGTAGCGCTTCACTGTGAAGAAAGCAAGATGCAATTTTAGAACTTTA 169
 Db |||||
 QY 99 GCTACAAATTTTAGCTCCCTGATTTGCTGTGGCAACAGCATGATCTTCGCGGAAGTG 158
 Db |||||
 QY 170 ACTACAGTTTTCCTGATTTGCTGTGGCAACAGCATGATCTTCGCGGAAGTG 229
 Db |||||
 QY 159 AGACGAGGCGCAATTTGAATTCCTCTTCAGAACATATGACAAAGCACACCTTTCCAGT 218
 Db |||||
 QY 230 AAACGAGGCGCAATTTGATGCTCTTTAGGACGTATGACAAAGGACATCACTTTTCAGT 289
 Db |||||
 QY 219 ATTTTAAGAGCTTCAAAACGTGTCCGATTAACCTTCAGCAACCCCTTATCTGAGCGGATG 278
 Db |||||
 QY 290 ATTTTAAGAGCTTCAAAACGTGTCCGATTAACCTTCAGCAACCCCTTATCTGAGCGGATG 349
 Db |||||
 QY 279 CAGGCTCGGCTGCACAGACCGAGTTCTCGGGGAAGAAATGAAGTTGTATTTTGCTC 338
 Db |||||
 QY 350 CCAGGCTCCAGCTGCATAGACTGAGTTTCTGGGAAGAAATGAAGTTATATTTTGCTC 409
 Db |||||
 QY 339 AGACTTTTACACATAGGAAGTTCCACCTGGCTCCGCC-AAATCCCGACAAACAGTTCCTCA 397
 Db |||||

Db 410 AGACTTTTACACATAGGAAGCTCACCTGGCTCCGCCAAATCCAGACAAAGCAGTTTCTGA 469
 QY 398 TCTCCCTCCCGCTCTCTCCCTGGTGGTGGAAACAAAGTAGAAGATGCCACCCCGTCA 457
 Db |||||
 QY 470 TCTCCCTCCCGCTCTCTCCCGAGTGGATGGAAACAAGTGGAGATCGGACCCAGTCA 529
 Db |||||
 QY 458 TAAATTTACGATCTTTTATATGCTCTCCAGCTGGGCGGACGAGAGAAAGTATGAATTCG 517
 Db |||||
 QY 530 TAAACTATGATCTCTTATATGCTCTCCAGCTGGGCGGACGAGAGAAAGTATGAATTCG 589
 Db |||||
 QY 518 ATGCAGCGACAGACACCACTCCAGTGTGGTGGTCCAGCTGTGTGAGAGTGACCAAGAGA 577
 Db |||||
 QY 590 ACGCAGCGACTGACACCACTCCAGCGTGTGGTCCATGTATGTGAGAGTGTCAAGAGA 649
 Db |||||
 QY 578 ATGAGGAGGAGAGGAA 594
 Db |||||
 QY 650 AGGAGGAAGAGAGGAA 666
 Db |||||

RESULT 13

PCT-US02-02176-1165

; Sequence 1165, Application PC/TUS0202176

; GENERAL INFORMATION:

; APPLICANT: Orr, Michael S.

; APPLICANT: Diggins, Michele

; APPLICANT: Zeng, Wen

; APPLICANT: Gene Logic, Inc.

; TITLE OF INVENTION: Gene Expression Profiles in Breast Tissue

; FILE REFERENCE: 44921-5070-WO

; CURRENT APPLICATION NUMBER: PCT/US02/02176

; CURRENT FILING DATE: 2002-01-25

; PRIOR APPLICATION NUMBER: US 60/263,757

; PRIOR FILING DATE: 2001-01-25

; PRIOR APPLICATION NUMBER: US 60/286,090

; PRIOR FILING DATE: 2001-04-25

; PRIOR APPLICATION NUMBER: US 60/292,517

; PRIOR FILING DATE: 2001-05-23

; NUMBER OF SEQ ID NOS: 1246

; SOFTWARE: Patent In Ver. 2.1

; SEQ ID NO 1165

; LENGTH: 2346

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: Genbank Accession No. U85267

PCT-US02-02176-1165

Query Match 69.1%; Score 413.8; DB 1; Length 2346;
 Best Local Similarity 85.1%; Pred. No. 4.8e-115;
 Matches 474; Conservative 0; Mismatches 82; Indels 1; Gaps 1;

QY 39 GAGCGAGTCGTTGTTAAGCGTCTGCCCGTGAAAGACGAGAAATGATTTTAGGACTTTA 98
 Db |||||
 QY 103 GACTGCTGGGTCTGTAGCGCTTTCACGTAAAGAAAGCAAGATGCAATTTTAGAACTTTA 162
 Db |||||
 QY 99 GCTACAAATTTTAGCTCCCTGATTTGCTGTGGCAACAGCATGATCTTCGCGGAAGTG 158
 Db |||||
 QY 163 ACTACAGTTTTCCTGATTTGCTGTGGCAACAGCATGATCTTCGCGGAAGTG 222
 Db |||||
 QY 159 AGACGAGGCGCAATTTGAATTCCTCTTCAGAACATATGACAAAGGACACCACTTTCCAGT 218
 Db |||||
 QY 223 AAACGAGGCGCAATTTGAGTCCCTCTTTAGGACGTATGACAAAGGACATCACTTTTCAGT 282
 Db |||||
 QY 219 ATTTTAAGAGCTTCAAAACGTGTCCGATTAACCTTCAGCAACCCCTTATCTGAGCGGATG 278
 Db |||||
 QY 283 ATTTTAAGAGCTTCAAAACGTGTCCGATTAACCTTCAGCAACCCCTTATCTGAGCGGATG 342
 Db |||||
 QY 279 CAGGCTCGGCTGCACAGACCGAGTTCTCGGGGAAGAAATGAAGTTGTATTTTGCTC 338
 Db |||||
 QY 343 CCAGGCTCCAGCTGCATAGACTGAGTTTCTGGGAAGAAATGAAGTTATATTTTGCTC 402
 Db |||||

Qy	339	AGACTTATACATATAGGAAGTTCACACCTGGCTCCGCC -AATCCGACAAACATTTCTCTCA	397
Db	403	AGACCTTACATATAGGAAGTTCACACCTGGCTCCGCCAATCCAGACAAAGCAATTTCTGA	462
Qy	398	TCTCCCTCCCGGCGCTCTCTCTCCGTTGGTGGAAAACAAGTAAAGATGCCACCCCGCTCA	457
Db	463	TCTCCCTCCCGCGCTCTCTCCGCAATGGGATGGAAAACAAGTGAAGATGCACCCCACTCA	522
Qy	458	TAAATTAGGATCTTTATATATGCATCTCCAACTGGGGCCAGGAGAAATATGAATGCTGC	517
Db	523	TAACTTAGATCTCTTATATGCCATCTCCAAAGCTGGGGCCAGGGGAAAAATATGAATTGC	582
Qy	518	ATGCAGCACAACAACCACTCCCACTGTGTGTGTCTCACCTGTGTGAAGTGAACCAAGAGA	577
Db	583	ACGCAGCACCAGACACCACTCCCAAGCGTGTGTCTCATGTATGTGAAGTGAATCAAGAGA	642
Qy	578	ATTGAGAGAAAGAGAA 594	
Db	643	AGGAGAGAAAGAGGAA 659	

```

RESULT 14
PCT-US02-25766-4321
Sequence 4321, Application PC/TUS0225766
GENERAL INFORMATION:
APPLICANT: GENE LOGIC, INC.
APPLICANT: MUNGER, William E
APPLICANT: FAULK, Ronald
APPLICANT: SUN, Hongwei
APPLICANT: SASAI, Hitoeohi
APPLICANT: WAGA, Iwao
APPLICANT: YAMAMOTO, Jun
TITLE OF INVENTION: Gene Expression Profiles in Glomerular Diseases
FILE REFERENCE: 44921-5068-WO
CURRENT APPLICATION NUMBER: PCT/US02/25766
CURRENT FILING DATE: 2002-08-14
PRIOR APPLICATION NUMBER: US 60/311,837
PRIOR FILING DATE: 2001-08-14
NUMBER OF SEQ ID NOS: 13946
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4321
LENGTH: 2346
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Genbank Accession No. U85267
PCT-US02-25766-4321

```

Query Match	69.1%;	Score 413.8;	DB 1;	Length 2346;
Best Local Similarity	85.1%;	Pred. No. 4.8e-15;		
Matches 474;	Conservative	0;	Mismatches 82;	Indels 1;
				Gaps 1

[illegible]

Accession	Sequence	Position
D8	AGACCTTACACATAGGAAAGCTCACACCTGGCTCCGCGCAAAATCCAGACAGAGATTTCTGA	462
OY	TCCTCCCTCCGGGCTCTCTCTCCGTTGGCTGGAAACAAGTGAAGATGCGCCCCCGCA	457
D8	TCCTCCCTCCGGGCTCTCGGCAATGGATGTAACCAAGTGAATGCGACCCCATCA	522
OY	TAAATTACATCTTTTATATATGCAATCTCCACGCTGGGGCAGAGAGAAATATGAATCC	517
D8	TAAACTTGTATCTCTTATATATGCAATCTCCACGCTGGGGCAGAGAGAAATATGAATTC	582
OY	ATGCAAGCAGACAGACACCACTCCCAAGTGTGATGCCACGTTGTGAGAGTGAACCAAGAGA	577
D8	ACCGAGGACATGCACACCACTCCCAAGGATGTGATCCATGTATGTAGAGATATCAAGAGA	642
OY	ATGAGAGAGAAAGAGAA 594	
D8	AGGAGAGAAAGAGAA 659	

```

RESULT 15
US-10-044-090-255
; Sequence 255, Application US/10044090
; GENERAL INFORMATION:
; APPLICANT: Oiga Bandman
; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
; FILE REFERENCE: PA-0028 US
; CURRENT APPLICATION NUMBER: US/10/044,090
; CURRENT FILING DATE: 2002-01-09
; NUMBER OF SEQ. ID NOS.: 850
; SOFTWARE: PERL Program
; SEQ ID NO 255
; LENGTH: 2355
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 5511889CB1
US-10-044-090-255

```

Query Match	69.1%;	Score 413.8;	DB 38;	Length 2355;
Best Local Similarity	85.1%;	Pred. No. 4.8e-115;		
Matches 474;	Conservative	0;	Mismatches 82;	Indels 1;
				Gaps 1;

Oy	39	GAGGAGGCGCTGGTTAAAGCGTCGCGCCGGTAAAAAAGCAGATGATTTTAGGACCTTTA	98
Db	110	GACTGCGGGGTCTGTAAGCGCTTCACTGTAAAGAAAGCAAGATGCATTTTAGAACTTTA	169
Oy	99	GCTACAAATTTTAGCTCCCTCATATTGCTGTGTGGCAACGATGATGCTTCAAGGAAAGTG	158
Db	170	ACTACAGTTTATGCTCTCCCTCATATTGCTGTGTGGCAACGATGATTCCTCAGGAAAGTG	229
Oy	159	AGACGAGGGCCCAATTTGAAATCCCTCTTCAGAACATATGACAAAGACACCACTTCCAGT	218
Db	230	AAACCGAGGCCCAATTTTGAGTCCCTCTTGAAGCATATGACAAGACATCACTTTCAGT	289
Oy	219	ATTTTAAAGCTTCAAACTGTGTCCGATTAACCTTCAGCAACCCCTTATCTCGACCGGATG	278
Db	290	ATTTTAAAGCTTCAAAAGATCGAGATTAACCTTCAGCAACCCCTTCTCGCAGCAGATG	349
Oy	279	CCAGGCTCGGCTGCACAAAGCCAGTTCTCTGGGGAAGAAATGAAGTTGTATTTTGCTC	338
Db	350	CCAGGCTCCAGCTGCATTAAGACTGAGTTTCTGGGAAAGAAATGAAGTTTATTTTGCTC	409
Oy	339	AGACTTTACATATAGGAAGTTCAACCTGGGTCCGCC- AATCCGCAAAACAGTTCCTCA	397
Db	410	AGACTTTACATATAGGAAGTTCAACCTGGGTCCGCCAAATCCAGCAACACAGTTTCTGA	469
Oy	398	TTTCCCTTCGGGCTCTCTCTCCGTTGGCTGGAAACAAGTAGAAGATGCCACCCCGTCA	457
Db	470	TTTCCCTTCGGGCTCTCTCTCCGCAAGTGGGATGGAAACAAGTAGAAGATGCCACCCCGTCA	529
Oy	458	TAAATTTAGATCTTTTATATGCATCTCCAAAGCTGGGGCCAGAGAGAAGTATGAAGTGC	517

```

Db 530 TAACTATGATCTTATATGCCATCTCCAAGCTGGGCGGAAAGTATGAATTGC 589
Qy 518 ATGCAGCGACAGACACCACTCCCAAGTGTGGTCCAGTGTGTGAGAGTGACCAAGAGA 577
Db 590 ACGCAGCGACTGACACCACTCCCAAGTGTGGTCCAGTGTGTGAGAGTGATCAAGAGA 649
Qy 578 ATGAGGAGGAGAGGAA 594
Db 650 AGGAGGAGGAGGAA 666

```

Search completed: December 14, 2002, 19:55:40
 Job time : 1047.77 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2002, 15:07:45 ; Search time 72.0383 Seconds
(without alignments)
9923.385 Million cell updates/sec

Title: US-09-782-953-11

Perfect score: 2331

Sequence: 1 tttttttttccccaggagt.....gggcacagtatgctcttac 2331

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*

1: /cgn2_6/ptodata/1/ina/5A COMB.seq:*

2: /cgn2_6/ptodata/1/ina/5B COMB.seq:*

3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*

4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*

5: /cgn2_6/ptodata/1/ina/6C COMB.seq:*

6: /cgn2_6/ptodata/1/ina/6D COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2059.2	88.3	2174	2	US-08-665-040-1
2	246.8	10.6	251	4	US-09-602-877A-83
3	52.8	2.3	7218	1	US-08-232-463-14
4	41	1.8	1770	4	US-08-939-309-7
5	40	1.7	40	2	US-08-665-040-9
6	39.6	1.7	2251	4	US-08-991-677-11
7	38.2	1.6	7218	1	US-08-232-463-14
8	38	1.6	327	4	US-09-889-595-7
9	38	1.6	5058	4	US-09-889-595-1
10	37.6	1.6	5194	1	US-09-599-653-1
11	37.6	1.6	5194	2	US-08-642-846-1
12	37.6	1.6	5194	4	US-09-264-604-1
13	37.2	1.6	6243	2	US-09-056-075-1
14	36.8	1.6	289	4	US-09-007-005-17
15	36.8	1.6	289	4	US-09-007-005-17
16	36.8	1.6	289	4	US-09-244-796-17
17	36.8	1.6	289	4	US-09-244-796-17
18	36.6	1.6	960	3	US-08-556-419-9
19	36.6	1.6	246240	2	US-08-724-394A-20
20	36.6	1.6	246240	2	US-08-724-394A-21
21	36.6	1.6	246240	2	US-08-724-394A-22
22	36.4	1.6	17656	4	US-09-433-579-3
23	36.2	1.6	583	4	US-09-404-671-5
24	36.2	1.6	583	4	US-09-404-671-7
25	36.2	1.6	162450	4	US-09-345-882-1
26	35.8	1.5	65042	4	US-09-784-316-3
27	35.4	1.5	4285	4	US-09-410-464-1

28	35.4	1.5	7425	4	US-09-453-702B-212
29	35.2	1.5	1032	4	US-08-482-073-8
30	35	1.5	87350	3	US-08-781-891-79
31	35	1.5	87343	4	US-09-791-211-3
32	35	1.5	246240	2	US-08-724-394A-20
33	35	1.5	246240	2	US-08-724-394A-21
34	35	1.5	246240	2	US-08-724-394A-22
35	34.2	1.5	688	1	US-08-139-937-9
36	34.2	1.5	688	5	PCT-US93-11310-9
37	34.2	1.5	5063	1	US-08-185-432-1
38	34.2	1.5	5093	1	US-08-468-036-23
39	34.2	1.5	5093	2	US-08-376-843-23
40	34	1.5	8703	4	US-08-961-527-177
41	34	1.5	19124	2	US-08-487-826B-13
42	33.8	1.5	667	4	US-09-123-912-99
43	33.8	1.5	667	4	US-09-643-597-99
44	33.8	1.5	2829	3	US-08-851-843A-53
45	33.8	1.5	2829	4	US-08-974-549A-220

ALIGNMENTS

RESULT 1

US-08-665-040-1
; Sequence 1, Application US/08665040
; Patent No. 5869318

GENERAL INFORMATION:

APPLICANT: ESTIVILL PALLEJA, XAVIER
APPLICANT: FUENTES, JUAN JOSE
APPLICANT: PRITCHARD, MELANIE
TITLE OF INVENTION: A NEW GENE SEQUENCE OF THE
TITLE OF INVENTION: DOWN SYNDROM CRITICAL REGION OF HUMAN CHROMOSOME 21,
TITLE OF INVENTION: IDENTIFIED BY A NEW "ALU-SPlicing PCR" TECHNIQUE,
TITLE OF INVENTION: CODING FOR A PROLINE-RICH PROTEIN (DSCR1) HIGHLY
TITLE OF INVENTION: EXPRESSED IN FOETEL BRAIN AND IN HEART AND METHOD
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: LADAS & PARRY
STREET: 26 WEST 61ST STREET
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10023

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORD PERFECT 5.1 FOR DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/665,040
FILING DATE: JUNE 7, 1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: ES P9501140
FILING DATE: JUNE 7, 1995
ATTORNEY/AGENT INFORMATION:
NAME: JANET I. CORD
REGISTRATION NUMBER: 33,778
REFERENCE/DOCKET NUMBER: U010815-9
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 708-1800
TELEFAX: (212) 246-8959

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 2174 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: cDNA for mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO

ORIGINAL SOURCE: human
ORGANISM: Homo sapiens
DEVELOPMENTAL STAGE: foetal
TISSUE TYPE: Brain
IMMEDIATE SOURCE: gene library of cDNA
LIBRARY: gene library of cDNA from foetal
CLONE: BC-17.8-1 and BC-17.8-2
POSITION IN GENOME:
CHROMOSOME/SEGMENT: chromosome 21/YAC 72H9
MAP POSITION: 21q22.1-q22.2
FEATURE:
NAME/KEY: cDNA for mRNA, BC-17.8
LOCATION: 1..2174
FEATURE:
NAME/KEY: untranslated 5'
LOCATION: 1..48
FEATURE:
NAME/KEY: coding sequence
LOCATION: 49..560
OTHER INFORMATION: Down Syndrome critical
FEATURE:
NAME/KEY: DSCR1
LOCATION: 1..171 PEPTIDES
IDENTIFICATION METHOD: translation of the
OTHER INFORMATION: - deduced protein
OTHER INFORMATION: - proline-rich protein domains
OTHER INFORMATION: - glutamic acid-rich protein domains
OTHER INFORMATION: - leucine/phenylalanine-rich protein domains
FEATURE:
NAME/KEY: untranslated 3'
LOCATION: 564..2174
FEATURE:
NAME/KEY: 2 poly (A)
LOCATION: 1541..1546 AND 2132..2137
US-08-665-040-1

Query Match 88.3%; Score 2059.2; DB 2; Length 2174;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 2001; Conservative 0; Mismatches 20; Indels 1; Gaps 1;

QY 231 GCCAAATTTAGTCCCTTTTAGAGCTATGACAGACATCACTTTCAGTATTTTAAAG 290
Db 58 GCCAAATTTAGTCCCTTTTAGAGCTATGACAGACATCACTTTCAGTATTTTAAAG 117
QY 291 AGCTTCAAAGAGTCAGATTAATTAATTGAGCAACCCCTTCCGAGAGATGCGAGGCTC 350
Db 118 AGCTTCAAAGAGTCAGATTAATTAATTGAGCAACCCCTTCCGAGAGATGCGAGGCTC 177
QY 351 CAGCTGATAGACTGAGTTCCTGGAAAGAAATGAAGTTATATTTTGTCTCAGACCTTA 410
Db 178 CAGCTGATAGACTGAGTTCCTGGAAAGAAATGAAGTTATATTTTGTCTCAGACCTTA 237
QY 411 CACATAGGAAGCTCACACCTGGCTCCGCCAAATCCAGACAGAGTTTCTGATCTCCCT 470
Db 238 CACATAGGAAGCTCACACCTGGCTCCGCCAAATCCAGACAGAGTTTCTGATCTCCCT 297
QY 471 CCCCCCTCTCCGAGTGGAGTGAACAAGTGAAGATGCAACCCAGTCATTAATAT 530
Db 298 CCCCCCTCTCCGAGTGGAGTGAACAAGTGAAGATGCAACCCAGTCATTAATAT 357
QY 531 GATCTCTTATATGCTATCTCCAGCTGGGGCCAGGGGAAAAGTATGATATGACGACGG 590
Db 358 GATCTCTTATATGCTATCTCCAGCTGGGGCCAGGGGAAAAGTATGATATGACGACGG 417
QY 591 ACTGACACCACTCCGAGGTGTGTGTCATGTATGTGAGTTGCAAGAGAAAGAGAA 650
Db 418 ACTGACACCACTCCGAGGTGTGTGTCATGTATGTGAGTTGCAAGAGAAAGAGAA 477
QY 651 GAAAGAGAAATGGAAGATGAGAGACCTAAGCCAAAATTTATCCAGACGAGAGCCG 710
Db 478 GAAAGAGAAATGGAAGATGAGAGACCTAAGCCAAAATTTATCCAGACGAGAGCCG 537

QY 711 GAGTACACGCCGATCCACTCAGCTGAACCTGGACCGGAGAGAGCATTCGAATCAT 770
Db 538 GAGTACACGCCGATCCACTCAGCTGAACCTGGACCGGAGAGAGCATTCGAATCAT 597
QY 771 ACTCAGGAGAGAAATCTTTTACTGTGAGAGTGTGCTGTCACGACTTCTTGGAGTGGCA 830
Db 598 ACTCAGGAGAGAAATCTTTTACTGTGAGAGTGTGCTGTCACGACTTCTTGGAGTGGCA 657
QY 831 GCCAGATTCGGGGTGGCAGAAATCCCAATTCATGTGCTCAGAAAGATCAAGCCCTG 890
Db 658 GCCAGATTCGGGGTGGCAGAAATCCCAATTCATGTGCTCAGAAAGATCAAGCCCTG 717
QY 891 TCCCTGTTTAAATGCTGCAACCACTTAATGTCATGACCCGGGAATGACTTGGG 950
Db 718 TCCCTGTTTAAATGCTGCAACCACTTAATGTCATGACCCGGGAATGACTTGGG 777
QY 951 CAATCACTGAGTTGTGTGATTCGACAAAGACATTTGGGACTGTCTTGAAAAACAGAT 1010
Db 778 CAATCACTGAGTTGTGTGATTCGACAAAGACATTTGGGACTGTCTTGAAAAACAGAT 837
QY 1011 AATGATAGTGTGTGATTCGTTCTTTCTGTGATGTTCTGTCTGTGCAAGGCGAGTT 1070
Db 838 AATGATAGTGTGTGATTCGTTCTTTCTGTGATGTTCTGTCTGTGCAAGGCGAGTT 897
QY 1071 GATCAGTACGCTCAGAGAGAGAGCTTCTGTGTTTCTAAGTGGCTGAGGGGCACTCTTA 1130
Db 898 GATCAGTACGCTCAGAGAGAGAGCTTCTGTGTTTCTAAGTGGCTGAGGGGCACTCTTA 957
QY 1131 CTGGTAGAAGAGGTACACACAGAAAGCCGCTAGTGCAGAGAGGTTGTGAACAGCAGC 1190
Db 958 CTGGTAGAAGAGGTACACACAGAAAGCCGCTAGTGCAGAGAGGTTGTGAACAGCAGC 1017
QY 1191 AATGCAATGTGAAATTTGATGAGCTTCTTTCTTCTCCCTCATGTTCTCATGTTGTGATG 1250
Db 1018 AATGCAATGTGAAATTTGATGAGCTTCTTTCTTCTCCCTCATGTTCTCATGTTGTGATG 1077
QY 1251 TATATTACGATTTTACAGAGCTTAACCTTGTGTAATTAAGTTACACCTGTGTGTT 1310
Db 1078 TATATTACGATTTTACAGAGCTTAACCTTGTGTAATTAAGTTACACCTGTGTGTT 1137
QY 1311 TACATCTTTTGGAGGCCAGGAAACGTTTGAAAACGTTATCAGCTTCCAGATTCTCG 1370
Db 1138 TACATCTTTTGGAGGCCAGGAAACGTTTGAAAACGTTATCAGCTTCCAGATTCTCG 1197
QY 1371 GATTCGACCTTTTGGACAGACACTTGTGCGGAATCTTCTCGGAATGCAATCACTC 1430
Db 1198 GATTCGACCTTTTGGACAGACACTTGTGCGGAATCTTCTCGGAATGCAATCACTC 1257
QY 1431 AGCATCCCAACCGTGCACAGTGAATGCTTGTGCAAAAGAAATGATCTGAAT 1490
Db 1258 AGCATCCCAACCGTGCACAGTGAATGCTTGTGCAAAAGAAATGATCTGAAT 1317
QY 1491 CCTGTGTAGAAATTTAGCTTATCAATTCAGAGATAGAGTTTCACTGCCAATTTAGT 1550
Db 1318 CCTGTGTAGAAATTTAGCTTATCAATTCAGAGATAGAGTTTCACTGCCAATTTAGT 1377
QY 1551 GGGTGAAGAAATTTAGTATAGTGTGTTGGATCGGACCTCACTTCTGTGTTCTTTA 1610
Db 1378 GGGTGAAGAAATTTAGTATAGTGTGTTGGATCGGACCTCACTTCTGTGTTCTTTA 1437
QY 1611 TGTGAGTGTGTTTATATCATGATCATAGCCAAAATTTTGGAAAATGTTGTTGAGA 1670
Db 1438 TGTGAGTGTGTTTATATCATGATCATAGCCAAAATTTTGGAAAATGTTGTTGAGA 1497
QY 1671 TAGTTGTTCTTTTACCCAGAGACATCAAGATACACTTGTAAATTAAGCTGATAGCA 1730
Db 1498 TAGTTGTTCTTTTACCCAGAGACATCAAGATACACTTGTAAATTAAGCTGATAGCA 1557
QY 1731 TATATTCAATCCTGTGTGACACTTGGTGAAGAAAGTATGAGCATGGAGCTAAGATGAT 1790
Db 1558 TATATTCAATCCTGTGTGACACTTGGTGAAGAAAGTATGAGCATGGAGCTAAGATGAT 1617
QY 1791 TAACTTAACCTGTGAATCATATGTTGTGAGAAAAGCTGTGTTCCCATGTCTAAGAGACTTGA 1850


```

US-08-939-309-7

Query Match          1.8%; Score 41; DB 4; Length 1770;
Best Local Similarity 51.4%; Pred. No. 0.028;
Matches 95; Conservative 0; Mismatches 90; Indels 0; Gaps 0.

QY  2011  CTTACAAATACGGTAACATGTTTGTCATTTTGAATACTTAATGCCAAGTAACAATGCAT 2070
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db   232  CTTTCACACGATTAACAAAAATTTACTTATTTAGCTGTACACAAAAGATTCACAAACAATATG 173
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY  2071  GCTTTGCAATTTGGAGAGATGGTTTATTTCTTTGAGAACCAATATGTTTGCAATTAATG 2130
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db   172  CCTTCACACATGTTGTACCATGGGGGATTTCTGAGGTAACCAATGTATGCTATTTTAAATT 113
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY  2131  CTTTATTTGTTCATATCAAGAAATGATTTGAACGTTCCCAAAACCGTGTACGTAAGT 2190
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db   112  CCTTATATGTTAGATCATTAACCTGGGCAAAAGTCGCTTTTCCCTTAGTAAATGAAATTG 53
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY  2191  GTAAG 2195
      |||  |||  |||
Db   52  GCATG 48
      |||  |||  |||

RESULT 5
US-08-665-040-9/c
; Sequence 9, Application US/08665040
; Patent No. 5869318
; GENERAL INFORMATION:
; APPLICANT: ESTIVILL, PALLEJA, XAVIER

```

```

1      RESULT 5
2      US-08-665-040-9/c
3      Sequence 9, Application US/08665040
4      Patent No. 5869318
5      GENERAL INFORMATION:
6      APPLICANT: ESTIVILL, PALIUEJA, XAVIER
7      APPLICANT: FUENTES, JUAN JOSE
8      APPLICANT: PRITCHARD, MELANIE
9      TITLE OF INVENTION: A NEW GENE SEQUENCE OF THE
10     TITLE OF INVENTION: DOWN SYNDROM CRITICAL REGION OF HUMAN CHROMOSOME 21
11     TITLE OF INVENTION: IDENTIFIED BY A NEW "ALU-SPlicing PCR" TECHNIQUE,
12     TITLE OF INVENTION: CODING FOR A PROLINE-RICH PROTEIN (OSCR1) HIGHLY
13     TITLE OF INVENTION: EXPRESSED IN FOETEL BRAIN AND IN HEART AND METHOD
14     NUMBER OF SEQUENCES: 12
15     FOR CHARACTERIZING IT.
16     CORRESPONDENCE ADDRESSES:
17     ADDRESSEE: LADAS & PARRY
18     STREET: 26 WEST 61ST STREET
19     CITY: NEW YORK
20     STATE: NEW YORK
21     COUNTRY: USA
22     ZIP: 10023
23     COMPUTER READABLE FORM:
24     MEDIUM TYPE: FLOPPY DISK
25     COMPUTER: IBM PC COMPATIBLE
26     OPERATING SYSTEM: PC-DOS/MS-DOS
27     SOFTWARE: WORD PERFECT 5.1 FOR DOS
28     CURRENT APPLICATION DATA:
29     APPLICATION NUMBER: US/08/665,040
30     FILING DATE: JUNE 7, 1996
31     CLASSIFICATION: 435
32     PRIOR APPLICATION DATA:
33     APPLICATION NUMBER: ES P9501140
34     FILING DATE: JUNE 7, 1995
35     ATTORNEY/AGENT INFORMATION:
36     NAME: JANET I. CORD
37     REGISTRATION NUMBER: 33,778
38     REFERENCE/DOCKET NUMBER: U010815-9
39     TELECOMMUNICATION INFORMATION:
40     TELEPHONE: (212) 708-1600
41     TELEFAX: (212) 246-8959
42     INFORMATION FOR SEQ ID NO: 9:
43     SEQUENCE CHARACTERISTICS:
44     LENGTH: 40 base pairs
45     TYPE: nucleic acid
46     STRANDEDNESS: single
47     TOPOLOGY: unknown
48     MOLECULE TYPE: DNA (genomic)
49     US-08-665-040-9

```

Query Match 1.7%; Score 40; DB 2; Length 40;
 Best Local Similarity 100.0%; Pred. No. 0.005;
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 240 GAGTCCCTTTAGGACGTATGACAAGGACATCACCTTTC 279
 DB 40 GAGTCCCTTTAGGACGTATGACAAGGACATCACCTTTC 1

RESULT 6

US-08-991-677-11/c
 ; Sequence 11, Application US/08991677A
 ; Patent No. 6252135

; GENERAL INFORMATION:
 ; APPLICANT: Chiang, Vincent L

; APPLICANT: Carraway, Daniel T
 ; APPLICANT: Smeltzer, Richard H

; TITLE OF INVENTION: Production of Syringyl Lignin in Gymnosperms
 ; FILE REFERENCE: 50617

; CURRENT APPLICATION NUMBER: US/08/991.677A
 ; CURRENT FILING DATE: 1997-12-16

; EARLIER APPLICATION NUMBER: US 60/033,381
 ; EARLIER FILING DATE: 1996-12-16

; NUMBER OF SEQ ID NOS: 11
 ; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 11
 ; LENGTH: 2251

; TYPE: DNA
 ; ORGANISM: Pinus taeda

US-08-991-677-11

Query Match

Best Local Similarity 1.7%; Score 39.6; DB 4; Length 2251;
 Matches 84; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 1952 TTGTGAATTTTCTAGTTGTAATGCTGCTGTAATAGGTATATATTTGGCC 2011
 DB 326 TGTTAGCTTTATTTAAAGTTTGTAGATTTGTATATAAAGTTGTTATTTGAAG 267

QY 2012 TTACAATACGTAACATGTTGTGCTATTTTGAATATCTTAATGCCAAGTAACAATGCATG 2071
 DB 266 TTTTAATAAATTTGTATAATATTTTTTTTAAATAGTTAAGTTATTTATTTAATC 207

QY 2072 CTTTGGAAATTTGGAATGCTTTTATCTTTGAGAAG 2109
 DB 206 TATTTAATTTGTTGAGAGGTTGTGATTGAGAAG 169

RESULT 7

US-08-232-463-14

; Sequence 14, Application US/08232463
 ; Patent No. 5670367

; GENERAL INFORMATION:
 ; APPLICANT: DORNER, F.

; APPLICANT: SCHEIFLINGER, F.
 ; APPLICANT: FALKNER, F. G.

; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
 ; NUMBER OF SEQUENCES: 52

; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Foley & Lardner

; STREET: 1800 Diagonal Road, Suite 500
 ; CITY: Alexandria

; STATE: VA
 ; COUNTRY: USA

; ZIP: 22313-0299

; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: IBM PC compatible
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/232,463

; FILING DATE:

; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/935,313
 ; FILING DATE:
 ; APPLICATION NUMBER: EP 91 114 300.6
 ; FILING DATE: 26-AUG-1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: BENT, Stephen A.
 ; REGISTRATION NUMBER: 29,768
 ; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (703)836-9300
 ; TELEFAX: (703)683-4109
 ; TELEX: 899149

; INFORMATION FOR SEQ ID NO: 14:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 7218 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; CLONE: PT2gpt-F1s

US-08-232-463-14

Query Match 1.6%; Score 38.2; DB 1; Length 7218;
 Best Local Similarity 8.8%; Pred. No. 0.5;
 Matches 25; Conservative 140; Mismatches 118; Indels 0; Gaps 0;

QY 1214 TTTCCTTTCTCCCTCATGTTCTCATGTTTGGCATGTATATTACTGATTACAGACTA 1273
 DB 1205 YY 1264

QY 1274 ACCTTGTTCGTATATAAGTTACACCTGTTGTTTACATCTTTGGGAGCCAGAA 1333
 DB 1265 YY 1324

QY 1334 AGCGTTTGGAAAACGTATACCTTTCCAGATTCTCGGATTCTCGACTCTTTCGCAACAGC 1393
 DB 1325 YY 1384

QY 1394 ACTTGCTTGGGAAGTCTCTCGGATGCAATCATCATGATCCCAACCTGCAACGTG 1453
 DB 1385 YY 1444

QY 1454 TAACCTGTGCTTTTGGAAAAGAGTTGATCTGAAATTCCTCTG 1496
 DB 1445 CTTCTATCTCTTAACCTACTTGCATAGTAGTAAATTACAGTG 1487

RESULT 8

US-09-889-595-7/c

; Sequence 7, Application US/09889595
 ; Patent No. 6410749

; GENERAL INFORMATION:
 ; APPLICANT: Aventis CropScience GmbH

; TITLE OF INVENTION: PROMOTERS FOR GENE EXPRESSION IN CARYOPSES OF PLANTS
 ; FILE REFERENCE: 514413-3885

; CURRENT APPLICATION NUMBER: US/09/889,595
 ; CURRENT FILING DATE: 2001-07-05

; PRIOR APPLICATION NUMBER: DE 100 32 379.0
 ; PRIOR FILING DATE: 2000-07-05

; NUMBER OF SEQ ID NOS: 18
 ; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 7
 ; LENGTH: 327

; TYPE: DNA
 ; ORGANISM: Triticum aestivum

US-09-889-595-7

Query Match 1.6%; Score 38; DB 4; Length 327;
 Best Local Similarity 47.5%; Pred. No. 0.079;
 Matches 113; Conservative 0; Mismatches 125; Indels 0; Gaps 0;


```
REFERENCE/DOCKET NUMBER: 960296.95238
TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-251-5000
TELEFAX: 608-251-9166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6243 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: misc feature
LOCATION: 3770..4013
OTHER INFORMATION: /note="RP4 origin of DNA transfer (orit) from
US-09-056-075-1
Query Match 1.6%; Score 37.2; DB 2; Length 6243;
Best Local Similarity 56.6%; Pred. No. 0.93;
Matches 69; Conservative 0; Mismatches 53; Indels 0; Gaps 0;
QY 2038 TTTTGAATACTTAATCCAGTACATGATGCTTGGAAATTTGGAAGATGCTTTTA 2097
DB 1499 TTTTCATATACGTAATATGACGTTTAATGTTTATGTTACTTCAATCATTAGAGTAA 1558
QY 2098 TTTTTCGACAGCAATATGTTGATTAAGCTTGTGATTCATATCAGAAATGTA 2157
DB 1559 TTTCTTATTTTAAAGCCTTTTCTTAAAGGCTTTTATTTTCTTAATACATTTA 1618
QY 2158 TT 2159
DB 1619 AT 1620
RESULT 14
US-09-007-005-17
Sequence 17, Application US/09007005B
Patent No. 6258558
GENERAL INFORMATION:
APPLICANT: Szostak, Jack W.
APPLICANT: Roberts, Richard W.
APPLICANT: Liu, Rih
TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
FILE REFERENCE: 00786/350003
CURRENT APPLICATION NUMBER: US/09/007,005B
EARLIER FILING DATE: 1998-01-14
EARLIER APPLICATION NUMBER: 60/035,963
EARLIER FILING DATE: 1997-01-27
EARLIER APPLICATION NUMBER: 60/064,491
NUMBER OF SEQ ID NOS: 33
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 17
LENGTH: 289
TYPE: RNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Translation template
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(289)
OTHER INFORMATION: n = A,T,C or G
US-09-007-005-17
Query Match 1.6%; Score 36.8; DB 4; Length 289;
Best Local Similarity 6.3%; Pred. No. 0.17;
Matches 15; Conservative 101; Mismatches 122; Indels 0; Gaps 0;
QY 615 GTCTGATATGTGAGTGTATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 674
DB 27 RURURARCRARARURURARCRARARURGRNRNRNRNRNRNRNRNRNRNRNRNR 86
```

```
QY 675 AGACCTAAGCCAAAATTTATCCAGACGAGCCGAGTACAGCCGATTCACCTCAGC 734
DB 87 RNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNR 146
QY 735 TGAACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 794
DB 147 RNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNR 206
QY 795 TGAGGTGCTGTGACGACCTTCTCGAGGTGACGACGACGACGACGACGACGACG 852
DB 207 RNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNR 264
RESULT 15
US-09-007-005-17/c
Sequence 17, Application US/09007005B
Patent No. 6258558
GENERAL INFORMATION:
APPLICANT: Szostak, Jack W.
APPLICANT: Roberts, Richard W.
APPLICANT: Liu, Rih
TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
FILE REFERENCE: 00786/350003
CURRENT APPLICATION NUMBER: US/09/007,005B
EARLIER FILING DATE: 1998-01-14
EARLIER APPLICATION NUMBER: 60/035,963
EARLIER FILING DATE: 1997-01-27
EARLIER APPLICATION NUMBER: 60/064,491
NUMBER OF SEQ ID NOS: 33
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 17
LENGTH: 289
TYPE: RNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Translation template
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(289)
OTHER INFORMATION: n = A,T,C or G
US-09-007-005-17
Query Match 1.6%; Score 36.8; DB 4; Length 289;
Best Local Similarity 10.4%; Pred. No. 0.17;
Matches 26; Conservative 94; Mismatches 130; Indels 0; Gaps 0;
QY 1277 TTTGTCGATATTAAGTTACCGTGTGTTTACATCTTTGGGAAGCAGGAAGC 1336
DB 265 TTTTAAAGCYGCVAYAGVAYGYTAAVCYGVAYGYTAAVCYGVAYGYTAAVCY 206
QY 1337 GTTTGAAAACGATACCTTCCAGATTCGGAATTCGACTCTTGGACACGACT 1396
DB 205 YNNNYSNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNN 146
QY 1397 TGCTTGGCACTCTTCGTAATGATTCAGTACGATCCCAACCGTGAACGTGTA 1456
DB 145 YNNNYSNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNN 86
QY 1457 CTGTGCTTTTGGAAAGAGTGAATTCCTGTGGAATTTAGCTTATACAT 1516
DB 85 YNNNYSNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNN 26
QY 1517 TCAGAGATA 1526
DB 25 YAYGYTAYTA 16
Search completed: December 14, 2002, 18:14:52
Job time : 123.038 secs
```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2002, 15:07:45 ; Search time 18.5118 Seconds
(without alignments)
9923.385 Million cell updates/sec

Title: US-09-782-953-1

Perfect score: 599

Sequence: 1 gaggtgcaaggaactccca.....gaggaggaaggaagagat 599

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*

1: /cgn2_6/ptodata/1/ina/5A COMB.seq:*

2: /cgn2_6/ptodata/1/ina/5B COMB.seq:*

3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*

4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*

5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq:*

6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	340.2	56.8	2174	2	US-08-665-040-1
2	36.6	6.1	3303	1	US-08-081-610-3
3	34	5.7	2118	4	US-09-221-017B-789
4	33	5.5	1685	1	US-09-640-305-1
5	33	5.5	1685	1	US-08-360-673-1
6	32.4	5.4	388	4	US-09-325-932A-108
7	32.2	5.4	3177	4	US-09-676-974-1
8	32	5.3	815	3	US-09-383-586-9
9	32	5.3	36519	3	US-08-923-137-2
10	31.8	5.3	4214	4	US-09-221-017B-293
11	31.6	5.3	1279	3	US-08-961-083-47
12	31.6	5.3	16995	4	US-08-961-527-82
13	31.2	5.2	4221	4	US-09-651-658-22
14	31.2	5.2	4221	4	US-09-650-855-22
15	31.2	5.2	16389	4	US-09-741-154-3
16	31	5.2	1930	2	US-08-737-716-1
17	31	5.2	1930	2	US-08-737-716-11
18	30.4	5.1	7431	4	US-08-665-040-9
19	30.4	5.1	7431	4	US-09-306-998-2
20	30.2	5.0	2518	4	US-09-433-699-3
21	30	5.0	1284	3	US-09-196-857-1
22	30	5.0	3489	2	US-08-728-323A-1
23	30	5.0	3489	4	US-09-298-568-1
24	30	5.0	32207	2	US-08-770-379-20
25	30	5.0	32207	4	US-08-757-669A-20
26	30	5.0	32207	4	US-09-230-371A-20
27	29.8	5.0	1512	2	US-08-909-965C-8

28	29.8	5.0	4031	2	US-08-993-118-1	Sequence 1, Appli
29	29.8	5.0	4031	3	US-08-845-528C-1	Sequence 1, Appli
30	29.8	5.0	4225	3	US-08-993-118-9	Sequence 9, Appli
31	29.8	5.0	4225	3	US-08-845-528C-9	Sequence 9, Appli
32	29.8	5.0	4265	4	US-09-061-709-1	Sequence 1, Appli
33	29.8	5.0	4415	4	US-09-486-580A-1	Sequence 1, Appli
34	29.8	5.0	12286	4	US-09-322-478-17	Sequence 17, Appli
35	29.8	5.0	12571	4	US-09-322-478-20	Sequence 20, Appli
36	29.6	4.9	2082	4	US-09-440-325A-2	Sequence 2, Appli
37	29.4	4.9	424	4	US-09-397-787-133	Sequence 133, App
38	29.4	4.9	1829	1	US-07-688-352C-15	Sequence 15, Appli
39	29.4	4.9	1829	2	US-08-474-379C-15	Sequence 15, Appli
40	29.4	4.9	1829	3	US-09-146-249A-15	Sequence 15, Appli
41	29.4	4.9	1829	3	US-08-206-188B-15	Sequence 15, Appli
42	29.4	4.9	1829	5	PCT-US91-02714-15	Sequence 15, Appli
43	29.4	4.9	2269	4	US-09-394-645-1	Sequence 1, Appli
44	29.4	4.9	2269	4	US-09-243-560B-1	Sequence 1, Appli
45	29.2	4.9	588	4	US-09-385-982-286	Sequence 286, App

ALIGNMENTS

RESULT 1

US-08-665-040-1
; Sequence 1, Application US/08665040
; Patent No. 5869318
; GENERAL INFORMATION:
; APPLICANT: ESTIVILL PALLEJA, XAVIER
; APPLICANT: FUENTES, JUAN JOSE
; APPLICANT: PRITCHARD, MELANIE
; TITLE OF INVENTION: A NEW GENE SEQUENCE OF THE
; TITLE OF INVENTION: DOWN SYNDROM CRITICAL REGION OF HUMAN CHROMOSOME 21,
; TITLE OF INVENTION: IDENTIFIED BY A NEW "ALU-SPLITTING PCR" TECHNIQUE,
; TITLE OF INVENTION: CODING FOR A PROLINE-RICH PROTEIN (DSCR1) HIGHLY
; TITLE OF INVENTION: EXPRESSED IN FOETAL BRAIN AND IN HEART AND METHOD
; TITLE OF INVENTION: FOR CHARACTERIZING IT.
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LADAS & PARRY
; STREET: 26 WEST 61ST STREET
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10023
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORD PERFECT 5.1 FOR DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/665,040
; FILING DATE: JUNE 7, 1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: ES P9501140
; FILING DATE: JUNE 7, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: JANET I. CORD
; REGISTRATION NUMBER: 33,778
; REFERENCE/DOCKET NUMBER: U010815-9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 708-1800
; TELEFAX: (212) 246-8959
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2174 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: cDNA for mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO

```

ORIGINAL SOURCE: human
ORGANISM: Homo sapiens
DEVELOPMENTAL STAGE: foetal
TISSUE TYPE: Brain
IMMEDIATE SOURCE: gene library of cDNA
LIBRARY: gene library of cDNA from foetal
CLONE: BC-17.8-1 and BC-17.8-2
POSITION IN GENOME:
CHROMOSOME/SEGMENT: chromosome 21/YAC 72H9
MAP POSITION: 21q22.1-q22.2
FEATURE:
NAME/KEY: cDNA for mRNA, BC-17.8
LOCATION: 1.2174
FEATURE:
NAME/KEY: untranslated 5'
LOCATION: 1.48
FEATURE:
NAME/KEY: coding sequence
LOCATION: 49.560
OTHER INFORMATION: Down Syndrome critical
region 1 (DSCR1)
FEATURE:
NAME/KEY: DSCR1
LOCATION: 1.171 PEPTIDES
IDENTIFICATION METHOD: translation of the
OTHER INFORMATION:
OTHER INFORMATION: - deduced protein
OTHER INFORMATION: - proline-rich protein domains
OTHER INFORMATION: - glutamic acid-rich protein domains
OTHER INFORMATION: - leucine/phenylalanine-rich protein domains
FEATURE:
NAME/KEY: untranslated 3'
LOCATION: 564.2174
FEATURE:
NAME/KEY: 2 poly (A)
LOCATION: 1541.1546 AND 2132.2137
US-08-665-040-1

```

```

Query Match          56.8%; Score 340.2; DB 2; Length 2174;
Best Local Similarity 88.6%; Pred. No. 5.4e-107;
Matches 380; Conservative 0; Mismatches 48; Indels 1; Gaps 1;

```

```

QY 167 GCCAAATTGATTCCTCTTCAAGACATATGACAGACACACCTTCCGATTTTAAAG 226
DB 58 GCCAAATTGATTCCTCTTCAAGACATATGACAGACACACCTTCCGATTTTAAAG 117
QY 227 AGCTTCAAACTGTCGCGATTAACCTTATCTGAGCGGATGCGAGCTG 286
DB 118 AGCTTCAAACTGTCGCGATTAACCTTATCTGAGCGGATGCGAGCTG 177
QY 287 CGGCTGCAAGACCGAGTTCCTGCGGAAAGAAATGAAGTTTGTCTCAGACTTGA 346
DB 178 CAGCTGATATGATGATGATTTCTGGAAGAAATGAAGTTTGTCTCAGACTTGA 237
QY 347 CACATAGGAAGTTGACACTTGGCTCCGCC -AATCCGACAAAGAGTTCCATCTCCCT 405
DB 238 CACATAGGAAGTTGACACTTGGCTCCGCCAATCCAGACAGAGCTTCTGATCTCCCT 297
QY 406 CCGGCTCTCTCCGCTGCGTGGGAAACAGTAGAAGATGCCACCCCGTCATTAATAC 465
DB 298 CCGGCTCTCTCCGCTGCGTGGGAAACAGTAGAAGATGCCACCCCGTCATTAATAC 357
QY 466 GATCTTTATATGCGCATCTCCAGCTGGGCGCAGAGAGAAATGAACTGACAGCGG 525
DB 358 GATCTTTATATGCGCATCTCCAGCTGGGCGCAGAGAGAAATGAACTGACAGCGG 417
QY 526 ACAGACACACTCCAGCTGGTGTGTGTCACAGTGTGAGAGTACCAAGATAGAGAG 585
DB 418 ACTGACACCACTCCAGCTGGTGTGTGTCACAGTGTGAGAGTATGATCAAGAGAGAG 477
QY 586 GAAAGAGAA 594
DB 478 GAAAGAGAA 486

```

```

RESULT 2
US-08-081-610-3
Sequence 3, Application US/08081610
Patent No. 5445941
GENERAL INFORMATION:
APPLICANT: Yang, Na N
TITLE OF INVENTION: Materials and Methods for Screening
TITLE OF INVENTION: Anti-Osteoporosis or Serum Lipid Lowering Agents
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSER: Allegretti and Witcoff, Ltd.
STREET: 10 S. Wacker Dr.
CITY: Chicago
STATE: IL
COUNTRY: U.S.A
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/081,610
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Heaphy, Barbara A
REGISTRATION NUMBER: 34,619
REFERENCE/DOCKET NUMBER: 93,402
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3303 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: mRNA
LOCATION: 2170..3303
FEATURE:
NAME/KEY: mRNA
LOCATION: 2214..3303
FEATURE:
NAME/KEY: mRNA
LOCATION: 2219..3303
FEATURE:
NAME/KEY: misc RNA
LOCATION: 3301..3303
OTHER INFORMATION: /note= "CDS Start, codon start = 1.
OTHER INFORMATION: translation M"
FEATURE:
NAME/KEY: TATA signal
LOCATION: 2170..2176
FEATURE:
NAME/KEY: misc feature
LOCATION: 1896..2306
OTHER INFORMATION: /note= "PB-301 -301 to +110"
FEATURE:
NAME/KEY: misc feature
LOCATION: 1976..2306
OTHER INFORMATION: /note= "PB-221 -221 to +110"
FEATURE:
NAME/KEY: misc feature
LOCATION: 2106..2306
OTHER INFORMATION: /note= "PB-91 -91 to +110"
FEATURE:
NAME/KEY: misc feature
LOCATION: 2137..2306
OTHER INFORMATION: /note= "PB-60 -60 to +110"

```

FEATURE:
NAME/KEY: misc_feature
LOCATION: 2150..2306
OTHER INFORMATION: /note= "pB-47 -47 to +110"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 2159..2306
OTHER INFORMATION: /note= "pB-38 -38 to +110"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 2159..2271
OTHER INFORMATION: /note= "TGPB-3 position -38 to
OTHER INFORMATION: +75"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 2159..2231
OTHER INFORMATION: /note= "TGPB-3 position -38 to
OTHER INFORMATION: +35"
US-08-081-610-3

Query Match 6.1%; Score 36.6; DB 1; Length 3303;
Best Local Similarity 60.6%; Pred. No. 0.032;
Matches 60; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
Qy 499 GGAGAGAGTATGAATGTCATGCGAGCGACACACACACTCCAGTGTGGTGTCCAGGTG 558
Db 2161 GGAGGAGGTATGAATTTTCAGCAGAGAGAAATAGAGAAAGCAGTGTGTGTCATGTGTG 2220
Qy 559 TGTGAGAGTGACCAAGAGAATGAGGAGGAGGAAGAGAG 597
Db 2221 TGTGTGTGAGAGAGAGGAGGAGGAGGAGGAGGAGAG 2259

RESULT 3

US-09-221-017B-789/c
; Sequence 789, Application US/09221017B
; Patent No. 6444799
; GENERAL INFORMATION:
; APPLICANT: Ross, Bruce C.
; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
; NUMBER OF SEQUENCES: 1120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/221,017B
; FILING DATE: 23-DEC-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP1182
; FILING DATE: 31-DEC-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP1546
; FILING DATE: 30-JAN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP2911
; FILING DATE: 09-APR-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU98/01023
; FILING DATE: 10-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Monroy, Gladys H.
; REGISTRATION NUMBER: 32,430
; REFERENCE/DOCKET NUMBER: 27340-20021.00

TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-813-5600
TELEFAX: 650-494-0792
TELEX: 706141
; INFORMATION FOR SEQ ID NO: 789:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2118 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: UNKNOWN
; ORIGINAL SOURCE:
; ORGANISM: PORYPHYROMONAS GINGIVALIS
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..2118
; US-09-221-017B-789
Query Match 5.7%; Score 34; DB 4; Length 2118;
Best Local Similarity 57.5%; Pred. No. 0.19;
Matches 61; Conservative 0; Mismatches 45; Indels 0; Gaps 0;
Qy 226 GAGCTTCAAACTGTCGCGATAAACTTCAGCAACCCCTTATCTGCAGCGGATGCCAGGCT 285
Db 1380 GAGTCTCTATATGTCGAGGAGGAAAATCCGTACCGCAGGTATGTAGACAATACAGGGGT 1321
Qy 286 GCGGTGCACAAGACCGAGTTCCTGGGAAGGAAATGAAGTTGTAT 331
Db 1320 TCGTCGCGAAGTGACGGAATCCTTCCGAGAAAATCAACTTTTTT 1275

RESULT 4

US-09-640-305-1/c
; Sequence 1, Application US/09640305
; Patent No. RE37447
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Yeh, Patrice
; TITLE OF INVENTION: MODIFIED KLUYVEROMYCES YEASTS, THEIR
; PREPARATION AND USE
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Rd. 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19002
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/640,305
; FILING DATE: 16-Aug-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/360,673
; FILING DATE: 06-FEB-1995
; APPLICATION NUMBER: WO PCT/FR93/00623
; FILING DATE: 23-JUN-1993
; APPLICATION NUMBER: FR 92/07785
; FILING DATE: 25-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, Julie K.
; REGISTRATION NUMBER: 38,619
; REFERENCE/DOCKET NUMBER: ST92040-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610)454-3839
; TELEFAX: (610)454-3808

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1685 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Kluyveromyces fragilis
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1683
OTHER INFORMATION: /product= "Protease B gene"
/gene= "K1.PR81"
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-640-305-1

Query Match 5.5%; Score 33; DB 1; Length 1685;
Best Local Similarity 53.5%; Pred. No. 0.37;
Matches 69; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

OY 88 TAGGAGCTTTAGTACATTTTATGCTCCCTGATGCTTGTGTGGCAAGATGATGCTT 147
DB 595 TAGAATCTTCTTCAACAAATTCACCAATGAGGCTTCTGTGATGAATGATGATCTT 536
OY 148 CAGCGAAGTGAGACGAGGCCAATTTGATCCCTCTTCAGAACATATGACAGACAC 207
DB 535 TTGTGAAGTAGCCACATTAACGAAAGGAAACCGCAATATCAAAAGTCTTGAAATAC 476
OY 208 CACCTTCCA 216
DB 475 CACCTTCCA 467

RESULT 5
US-08-360-673-1/c
Sequence 1, Application US/08360673
Patent No. 5679544
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
APPLICANT: Fournier, Alain
TITLE OF INVENTION: MODIFIED KLUYVEROMYCES YEASTS, THEIR
TITLE OF INVENTION: PREPARATION AND USE
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Rd. 3c43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19002
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/360,673
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR93/00623
FILING DATE: 23-JUN-1993
PRIOR APPLICATION DATA: FR 92/07785
APPLICATION NUMBER: FR 92/07785
FILING DATE: 25-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Julie K.
REGISTRATION NUMBER: 38,619
REFERENCE/DOCKET NUMBER: ST92040-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610)454-3839

TELEFAX: (610)454-3808
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1685 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Kluyveromyces fragilis
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1683
OTHER INFORMATION: /product= "Protease B gene"
/gene= "K1.PR81"
US-08-360-673-1

Query Match 5.5%; Score 33; DB 1; Length 1685;
Best Local Similarity 53.5%; Pred. No. 0.37;
Matches 69; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

OY 88 TAGGAGCTTTAGTACATTTTATGCTCCCTGATGCTTGTGTGGCAAGATGATGCTT 147
DB 595 TAGAATCTTCTTCAACAAATTCACCAATGAGGCTTCTGTGATGAATGATGATCTT 536
OY 148 CAGCGAAGTGAGACGAGGCCAATTTGATCCCTCTTCAGAACATATGACAGACAC 207
DB 535 TTGTGAAGTAGCCACATTAACGAAAGGAAACCGCAATATCAAAAGTCTTGAAATAC 476
OY 208 CACCTTCCA 216
DB 475 CACCTTCCA 467

RESULT 6
US-09-325-932A-108/c
Sequence 108, Application US/09325932A
Patent No. 6451604
GENERAL INFORMATION:
APPLICANT: Flinn, Barry
APPLICANT: Lasham, Annette
TITLE OF INVENTION: Compositions affecting programmed cell
TITLE OF INVENTION: death and their use in the modification of forestry plant develo
FILE REFERENCE: 1022
CURRENT APPLICATION NUMBER: US/09/325,932A
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 206
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO: 108
LENGTH: 388
TYPE: DNA
ORGANISM: Pinus radiata
US-09-325-932A-108

Query Match 5.4%; Score 32.4; DB 4; Length 388;
Best Local Similarity 50.6%; Pred. No. 0.25;
Matches 78; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

OY 364 CCTGGCTCCGCAATCCGACAAAGATTCATCTCCCTCCGCGCTTCTCCCGTT 423
DB 379 CCGAGCTTGGCAGGTAGATGATCTGCTTCACTAAGATCCGCAATATCTATGAT 320
OY 424 GCGTGAACAAAGTAGAATGCCACCCCGTCATTAATTCATGATCTTTATGCGCATC 483
DB 319 AATATGAAGGCGGTGTCCATAGCCCTCGTTGCAACCTGCGTTCTTTATCGCAATC 260
OY 484 TCCAGCTGGGCGAGAGAGAGATGAACTGC 517
DB 259 TACCACTGTGCTCTGATTAAGAAACCGCTTC 226

RESULT 7
US-09-676-974-1/c

; Sequence 1, Application US/09676974
; Patent No. 6391631
; GENERAL INFORMATION:
; APPLICANT: HACKER, JORG
; APPLICANT: SONNEN-BORN, ULRICH
; APPLICANT: SCHULZE, JURGEN
; APPLICANT: BLUM-OEHLER, GABRIELE
; APPLICANT: MALINKA, JURGEN
; APPLICANT: PROPERT, HANS
; TITLE OF INVENTION: BACTERIAL PLASMIDS
; FILE REFERENCE: 11347/268416/BET
; CURRENT APPLICATION NUMBER: US/09/676,974
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: PCT/EP98/01720
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: DE 197 13543.9
; PRIOR FILING DATE: 1997-04-02
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 3177
; TYPE: DNA
; ORGANISM: Escherichia coli
; US-09-676-974-1

Query Match 5.4%; Score 32.2; DB 4; Length 3177;
Best Local Similarity 59.1%; Pred. No. 1;
Matches 55; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

Qy 69 TGAAGAGCAGAGATGATTTAGGACCTTTAGTACATTTTAGCTCCCTGATGCTTG 128
Db 2938 TCAAAATGCTCAATGAATCATCGATTTTATCGAGTGTGTTTCAGATATCTGGATG 2879

Qy 129 TGGCAACAGATGATGCTTTCAGCGAAGTGAGA 161
Db 2878 TAATAAGGTTGAAGGTCTTCAGAGAATGAGACA 2846

RESULT 8
US-09-383-586-9
; Sequence 9, Application US/09383586
; Patent No. 6242419
; GENERAL INFORMATION:
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Abernethy, Nevin
; APPLICANT: Onrust, Rene
; APPLICANT: Kumble, Anand
; APPLICANT: Murison, Greg
; TITLE OF INVENTION: Compounds isolated from stromal cells
; FILE REFERENCE: 11000.1037c1
; CURRENT APPLICATION NUMBER: US/09/383,586
; CURRENT FILING DATE: 1999-08-26
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 815
; TYPE: DNA
; ORGANISM: Mouse
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (119)...(682)
; NAME/KEY: sig peptide
; LOCATION: (119)...(205)
; US-09-383-586-9

Query Match 5.3%; Score 32; DB 4; Length 815;
Best Local Similarity 54.2%; Pred. No. 0.53;
Matches 65; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

Qy 141 ATGCTTCAGCGAAGTGAGACAGCGGCGCAATTTGAATCCCTCTTCAGACATATGACA 200
Db 141 ATGCTTCAGCGAAGTGAGACAGCGGCGCAATTTGAATCCCTCTTCAGACATATGACA 200

Db 431 AAGTCTACACCCAAATGCAAGTGTCTTCTCCAAACTTACCACACGTCAGCATCCATGACA 490
Qy 201 AGGACACACCTTCCAGTATTTTAAAGAGCTTCAACGCTGTCGGGATAAACTTTCAGAACCC 260
Db 491 ACCCAGCCCCACAGTAGTTTATTGACAACTGTAACGGTTTTCAGCAACTACTCATCCACC 550

RESULT 9
US-08-923-137-2
; Sequence 2, Application US/08923137
; Patent No. 6083716
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Farina, Steven F.
; APPLICANT: Fisher, Krishna J.
; TITLE OF INVENTION: Chimpanzee Adenovirus Vectors
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr., P.O. Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: United States of America
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/923,137
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/024,700
; FILING DATE: 06-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: GNVN.021CIP1USA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 36519 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; US-08-923-137-2

Query Match 5.3%; Score 32; DB 3; Length 36519;
Best Local Similarity 56.7%; Pred. No. 5.4;
Matches 59; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

Qy 494 GCCCAGGAGAGAGTATGAACTGCATGCAGCGACAGACACACCTCCAGCTGTGTGTGTC 553
Db 25600 GTCAGGAGAGGAGGAGGAGATGCGAAGACTGGGACAGCACTCAGGAGGAGGAGGACA 25659

Qy 554 ACCTGTGTGAGAGTGACCAAGAGATGAGGAGGAAGAGGAGAG 597
Db 25660 GCCTGCAAGACAGTCTGGAGGAGACGAGGAGGAGGAGGAG 25703

RESULT 10
US-09-221-017B-293/c
; Sequence 293, Application US/09221017B
; Patent No. 6444799
; GENERAL INFORMATION:
; APPLICANT: Ross, Bruce C.
; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
; NUMBER OF SEQUENCES: 1120

```

CORRESPONDENCE ADDRESS:
ADDRESS: MORRISON & FOERSTER
STREET: 755 PAGE MILL ROAD
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FASTSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/221.017B
FILING DATE: 23-DEC-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P1182
FILING DATE: 31-DEC-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P1546
FILING DATE: 30-JAN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P22911
FILING DATE: 09-APR-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU98/01023
FILING DATE: 10-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: Montoy, Gladys H
REGISTRATION NUMBER: 32,430
REFERENCE/DOCKET NUMBER: 27340-20021.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-813-5600
TELEFAX: 650-494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 293:
SEQUENCE CHARACTERISTICS:
LENGTH: 4214 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: UNKNOWN
ORIGINAL SOURCE:
ORGANISM: PORPHYROMONAS GINGIVALIS
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1...4214
US-09-221-017B-293

Query Match      5.3%; Score 31.8; DB 4; Length 4214;
Best Local Similarity 59.3%; Pred. No. 1.7;
Matches 54; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 30 TGACTGAGAGCGAGCTGTTGCTTAAAGCGTCTGCCCCGTAAGAAAACAGAAATGATTTTA 89
DB 3420 TGCCTGGAAGAGCAAGATGATGCTGAGCAAGCTGTCCCATATTCAGGCCAGCTGATTGCA 3361
QY 90 GCGACTTTAGCTACAAATTTTAAAGCTCCCTGAT 120
DB 3360 TGAAGCTTAGCTCCCATGCTTCTCGAAGCAT 3330

RESULT 11
US-08-961-083-47/c
Sequence 47, Application US/08961083
Patent No. 6159469
GENERAL INFORMATION:
APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines
NUMBER OF SEQUENCES: 452
```

```

CORRESPONDENCE ADDRESS:
ADDRESS: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961.083
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 1279 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-961-083-47

Query Match      5.3%; Score 31.6; DB 3; Length 1279;
Best Local Similarity 49.4%; Pred. No. 0.96;
Matches 82; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 53 TTAAGCGTCTGCCCGGTAAAGAGAAATGATTTTAAAGACTTTAGTACAAATTTTAC 112
DB 733 TTTAGCTTTTACTTCTGCGATATGTTGTCACGAATTTTACTTCTTACCTGCAAGCTCTTC 674
QY 113 TCCCGATGTCCTTGTGTGCAAGCATGATGCTTTCAGCGAAAGTGAGACGAGGCCAAA 172
DB 673 TCCCTGTAGTCTTCTGGAATGTTACATTAACATCAACGATTTCCAGCTGATGACC 614
QY 173 TTTGAATCCCTCTTCAGAACATATGACAAGACACCACTTCCAGT 218
DB 613 TACCAATGGTCTTCGAAACCGAGGATGATGACCTGAACCAAGT 568

RESULT 12
US-08-961-527-82/c
Sequence 82, Application US/08961527
Patent No. 6420135
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
NUMBER OF SEQUENCES: 391
CORRESPONDENCE ADDRESS:
ADDRESS: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961.527
```

? FILING DATE:
 ? CLASSIFICATION: 424
 ? PRIOR APPLICATION DATA:
 ? APPLICATION NUMBER:
 ? FILING DATE:
 ? ATTORNEY/AGENT INFORMATION:
 ? NAME: Brookes, A. Anders
 ? REGISTRATION NUMBER: 36,373
 ? REFERENCE/DOCKET NUMBER: PB340P1
 ? TELECOMMUNICATION INFORMATION:
 ? TELEPHONE: (301) 309-8504
 ? TELEFAX: (301) 309-8512
 ? INFORMATION FOR SEQ ID NO: 82:
 ? SEQUENCE CHARACTERISTICS:
 ? LENGTH: 16995 base pairs
 ? TYPE: nucleic acid
 ? STRANDEDNESS: double
 ? TOPOLOGY: linear
 ? US-08-961-527-82

Query Match	5.3%	Score 31.6;	DB 4;	Length 16995;
Best Local Similarity	49.4%;	Pred. No. 4.6;		
Matches 82;	Conservative 0;	Mismatches 84;	Indels 0;	Gaps 0;
QY 53	TTAAGCGTCTGCCCGTGGAAAGCAGAAATGATTTTAGGACCTTTAGCTACAATTTTACG 112			
Db 7574	TTTAGCTTTTCTCGTGGATGTTGCACAAATTTAGTCTTTTACCTGCAAGGTCTTC 7515			
QY 113	TCCTTGATGCTTGTGTGGCAACGATGATCTTTCAGCGAAAGTAGAGACCGAGGCCAA 172			
Db 7514	TGCTTGGTAGTCTTCTGGGAATGTTACGATACATCAACGGTTCGCCAGCTAGTGACC 7455			
QY 173	TTTGAATCCCTCTTCAGAACATATGACAGGACACCACTTCCAGT 218			
Db 7454	TACCAATGGTCTTCGAAACAGGGATGAAATTGACCTTGAACCAAGT 7409			

RESULT 13
US-09-651-656-22
; Sequence 22, Application US/09651656
; Patent No. 6340566
; GENERAL INFORMATION:
; APPLICANT: MCCUTHEN-MALONEY, SANDRA
; APPLICANT: LAWRENCE LIVERMORE NATIONAL LABORATORY
; TITLE OF INVENTION: DETECTION AND QUANTITATION OF SINGLE NUCLEOTIDE
; TITLE OF INVENTION: POLYMORPHISMS, DNA SEQUENCE VARIATIONS, DNA MUTATIONS,
; TITLE OF INVENTION: DNA DAMAGE AND DNA MISMATCHES
; FILE REFERENCE: IL-10689
; CURRENT APPLICATION NUMBER: US/09/651.656
; CURRENT FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/192,764
; PRIOR FILING DATE: 2000-03-28
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 4221
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-651-656-22

	Query Match	5.2%	Score 31.2	DB 4	Length 4221	
	Best Local Similarity	54.3%	Pred. No. 2.7			
	Matches 63	Conservative 0	Mismatches 53	Indels 0	Gaps 0	
QY	483	CTCCAAGCTGGGGCCAGAGAGAAGTATGAATCTGCATSCAGCGACAGACACCACCTCCCG	542			
Db	2593	CTTCATTTGGSGCAGACAAATGGTGTATCTCTGAAGAGCGACCCCAACGAGGTTCGAT	2652			
QY	543	TGTGGTGGTCCACGTGTGTGAGAGTGACCAAGAGAATGACGAGCGAAGAGAGAGA	598			
Db	2653	TGTGGTCTTTCAGTATTCGGATTGGAAATACAGACTGTGGCTGAAAGAGCAAGTGA	2708			

Search completed: December 14, 2002, 18:13:40
Job time : 65.5118 secs

```

RESULT 14
US-09-650-855-22
; Sequence 22, Application US/09650855
; Patent No. 6365355
; GENERAL INFORMATION:
; APPLICANT: MCCUTHEN-MALONEY, SANDRA
; APPLICANT: LAWRENCE LIVERMORE NATIONAL LABORATORY
; TITLE OF INVENTION: CHIMERIC PROTEINS FOR DETECTION AND QUANTITATION OF DNA
; TITLE OF INVENTION: MUTATIONS, DNA SEQUENCE VARIATIONS, DNA DAMAGE AND DNA
; TITLE OF INVENTION: MISMATCHES
; FILE REFERENCE: IL-10284
; CURRENT APPLICATION NUMBER: US/09/650,855
; CURRENT FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/192,764
; PRIOR FILING DATE: 2000-03-28
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 4221
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-650-855-22

```

	Query Match	5.2%	Score 31.2;	DB 4;	Length 4221;
	Best Local Similarity	54.3%;	Pred. No. 2.7;		
	Matches 63;	Conservative 0;	Mismatches 53;	Indels 0;	Gaps 0;
QY	483	CTCCAAGCTGGGCGCAGGAGAAATGAAC	TGCATG	CAGCGACAGACACC	ACTCCCAAG 542
Db	2593	CTTCATTGTGGGCACAGCAAT	TGGTGTTATCCTGAAG	GCGCAGGCCAAGCAGG	GTGCGCAT 2652
QY	543	TGTGGTGGTCCACGTGTGTGAGACTG	ACCAAGAAATGAGGAGGA	GAGGGAAGAAGA	598
Db	2653	TGTGGTCTTTTCAGTATTCGGAT	TGGAAATACTACACACTGTGGCT	CTCAAAGGAAGTGA	2708

```

RESULT 15
US-09-741-154-3
; Sequence 3, Application US/09741154
; Patent No. 6437110
; GENERAL INFORMATION:
; APPLICANT: BEASLEY, Ellen M. et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001061
; CURRENT APPLICATION NUMBER: US/09/741,154
; CURRENT FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 16389
; TYPE: DNA
; ORGANISM: Human
US-09-741-154-3

```

	Query Match	5.2%	Score 31.2	DB 4	Length 16389	
	Best Local Similarity	54.3%	pred. No. 6.2			
	Matches 63	Conservative 0	Mismatches 53	Indels 0	Gaps 0	
QY	197	GACAAGGACACACCTTCCAGTATTTTAAAGAGCTTCAAACGCTGTCGCGATAAACTTCAGC	256			
Db	11592	GAGAGTGAGACCCCTGTATCAATATTTTAAAAATAAAAAATGTTGCGCGGCACATGGCTC	11651			
QY	257	AACCCCTTATCTCAGCCGATGCCAGCTCGGCTGCACAAGACCGAGTTCCTGGG	312			
Db	11652	ATGCCTGTATCCAGCACATTTGGAGAGCTGAGCGCGCGATCACAAGGTCATCAG	11707			

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2002, 15:07:45 ; Search time 18.45 Seconds
(without alignments)
9923.385 Million cell updates/sec

Title: US-09-782-953-5

Perfect score: 597

Sequence: 1 atggatttgaaggacttag.....caccgacccttagctga 597

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*

1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*

2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*

3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*

4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*

5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq.*

6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	398	66.7	2174	2	US-08-665-040-1 Sequence 1, Appl
2	38.4	6.4	3303	1	US-08-081-610-3 Sequence 3, Appl
3	37.8	6.3	7218	1	US-08-232-463-14 Sequence 14, Appl
4	37.6	6.3	7130	4	US-09-056-105-31 Sequence 31, Appl
5	36.2	6.1	6328	4	US-08-913-832A-1 Sequence 1, Appl
6	36.2	6.1	6328	4	US-09-249-181A-1 Sequence 1, Appl
7	34	5.7	1296	1	US-07-816-283-9 Sequence 9, Appl
8	34	5.7	1296	1	US-08-417-103-9 Sequence 9, Appl
9	34	5.7	2118	4	US-09-221-017B-789 Sequence 789, App
10	33.2	5.6	2277	1	US-08-676-967-2 Sequence 2, Appl
11	33.2	5.6	2277	1	US-08-676-967-2 Sequence 2, Appl
12	33.2	5.6	2277	2	US-09-098-487-2 Sequence 2, Appl
13	33.2	5.6	2289	4	US-09-312-038-3 Sequence 3, Appl
14	33	5.5	1685	1	US-09-840-305-1 Sequence 1, Appl
15	33	5.5	1685	1	US-08-360-673-1 Sequence 1, Appl
16	33	5.5	2089	1	US-08-552-142A-1 Sequence 1, Appl
17	33	5.5	2089	1	US-08-910-973-1 Sequence 1, Appl
18	33	5.5	2089	1	US-09-499-227-1 Sequence 1, Appl
19	33	5.5	2089	5	PCT-US95-05741-1 Sequence 1, Appl
20	33	5.5	4221	4	US-09-651-656-22 Sequence 22, Appl
21	33	5.5	4221	4	US-09-650-855-22 Sequence 22, Appl
22	32.4	5.4	435	4	US-09-397-787-282 Sequence 20, Appl
23	32.2	5.4	815	4	US-09-383-586-9 Sequence 282, App
24	32	5.4	16442	3	US-08-781-891-208 Sequence 9, Appl
25	32	5.4	3489	2	US-08-728-323A-1 Sequence 208, App
26	31.8	5.3	3489	4	US-09-298-568-1 Sequence 1, Appl
27	31.8	5.3	3489	4	US-09-298-568-1 Sequence 1, Appl

C 28	31.8	5.3	32207	2	US-08-770-379-20	Sequence 20, Appl
C 29	31.8	5.3	32207	4	US-08-757-669A-20	Sequence 20, Appl
C 30	31.8	5.3	32207	4	US-09-230-371A-20	Sequence 20, Appl
C 31	31.6	5.3	7015	4	US-09-177-249-6	Sequence 6, Appl
32	31.6	5.3	36519	3	US-08-923-137-2	Sequence 2, Appl
33	31.2	5.2	16389	4	US-09-741-154-3	Sequence 3, Appl
34	31	5.2	1462	1	US-08-552-142A-16	Sequence 16, Appl
35	31	5.2	1930	2	US-08-737-716-1	Sequence 1, Appl
36	31	5.2	1930	2	US-08-737-716-11	Sequence 11, Appl
37	31	5.2	1951	1	US-08-910-973-16	Sequence 16, Appl
38	31	5.2	1951	4	US-09-499-227-16	Sequence 16, Appl
39	31	5.2	15378	3	US-08-785-420-1	Sequence 1, Appl
C 40	30.8	5.2	873	4	US-09-134-001C-2206	Sequence 2206, App
C 41	30.8	5.2	28720	4	US-09-341-587-7	Sequence 7, Appl
C 42	30.6	5.1	3728	1	US-08-111-939-1	Sequence 1, Appl
C 43	30.6	5.1	51259	3	US-08-781-891-209	Sequence 209, App
C 44	30.4	5.1	40	2	US-08-665-040-9	Sequence 9, Appl
C 45	30.4	5.1	289	4	US-09-007-005-17	Sequence 17, Appl

ALIGNMENTS

RESULT 1

US-08-665-040-1

; Sequence 1, Application US/08665040

; Patent No. 5869318

; GENERAL INFORMATION:

; APPLICANT: ESTIVILL PALLEJA, XAVIER

; APPLICANT: FUENTES, JUAN JOSE

; APPLICANT: PRITCHARD, MELANIE

; TITLE OF INVENTION: A NEW GENE SEQUENCE OF THE

; TITLE OF INVENTION: DOWN SYNDROM CRITICAL REGION OF HUMAN CHROMOSOME 21,

; TITLE OF INVENTION: IDENTIFIED BY A NEW "ALU-SPLICING PCR" TECHNIQUE,

; TITLE OF INVENTION: CODING FOR A PROLINE-RICH PROTEIN (DSCR1) HIGHLY

; TITLE OF INVENTION: EXPRESSED IN FOETEL BRAIN AND IN HEART AND METHOD

; NUMBER OF SEQUENCES: 12

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: LADAS & PARRY

; STREET: 26 WEST 61ST STREET

; CITY: NEW YORK

; STATE: NEW YORK

; COUNTRY: USA

; ZIP: 10023

; COMPUTER READABLE FORM:

; MEDIUM TYPE: FLOPPY DISK

; COMPUTER: IBM PC COMPATIBLE

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: WORD PERFECT 5.1 FOR DOS

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/665,040

; FILING DATE: JUNE 7, 1996

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: ES P9501140

; FILING DATE: JUNE 7, 1995

; ATTORNEY/AGENT INFORMATION:

; NAME: JANET I. CORD

; REGISTRATION NUMBER: 33,778

; REFERENCE/DOCKET NUMBER: U010815-9

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 708-1800

; TELEFAX: (212) 246-8959

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2174 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: Double

; TOPOLOGY: circular

; MOLECULE TYPE: cdna for mRNA

; HYPOTHETICAL: No

; ANTI-SENSE: No

ORIGINAL SOURCE: human
 ORGANISM: Homo sapiens
 DEVELOPMENTAL STAGE: foetal
 TISSUE TYPE: Brain
 IMMEDIATE SOURCE: gene library of cDNA
 LIBRARY: gene library of cDNA from foetal
 CLONE: BC-17.8-1 and BC-17.8-2
 POSITION IN GENOME:
 CHROMOSOME/SEGMENT: chromosome 21/YAC 72H9
 MAP POSITION: 21q22.1-q22.2
 FEATURE:
 NAME/KEY: cDNA for mRNA, BC-17.8
 LOCATION: 1..2174
 FEATURE:
 NAME/KEY: untranslated 5'
 LOCATION: 1..48
 FEATURE:
 NAME/KEY: coding sequence
 LOCATION: 49..560
 OTHER INFORMATION: Down Syndrome critical
 FEATURE:
 NAME/KEY: DSCR1
 LOCATION: 1..171 PEPTIDES
 IDENTIFICATION METHOD: translation of the
 OTHER INFORMATION: DSCR1 sequence.
 OTHER INFORMATION: - deduced protein
 OTHER INFORMATION: - proline-rich protein domains
 OTHER INFORMATION: - glutamic acid-rich protein domains
 OTHER INFORMATION: - leucine/phenylalanine-rich protein domains
 FEATURE:
 NAME/KEY: untranslated 3'
 LOCATION: 564..2174
 FEATURE:
 NAME/KEY: 2 poly (A)
 LOCATION: 1541..1546 AND 2132..2137
 US-08-665-040-1

Query Match 66.7%; Score 398; DB 2; Length 2174;
 Best Local Similarity 87.6%; Pred. No. 4,2e-122;
 Matches 447; Conservative 0; Mismatches 60; Indels 3; Gaps 1;

88 GCCAAATTTGAATCCCTCTTCAGACATATGACAGACACCACTTCCAGTATTTTAAAG 147
 |||||
 58 GCCAAATTTGAGTCCCTCTTAGACGTAATGACAGACATCACTTTAGTATTTTAAAG 117
 |||||
 148 AGCTTCAAAGTGTCCGGATAAACTTCAGCAACCCCTTATCTGACGCCAGTCTG 207
 |||||
 118 AGCTTCAAAGTGTCCGGATAAACTTCAGCAACCCCTTCTCCGACAGATGCCAGCTC 177
 |||||
 208 CGGCTGCAACAGCCGAGTCTCTGGGGAGGAATGAAGTTGATTTGCTCAGACTTAA 267
 |||||
 178 CAGCTGATAGACTGAGTTCTGGGAAAGGAATGAAGTTATTTGCTCAGACTTAA 237
 |||||
 268 CACATAGGAAGTTCACACTGGCTCCGCCCAATCCGACAAAGTTCCATCTCCCT 327
 |||||
 238 CACATAGGAAGTTCACACTGGCTCCGCCCAATCCGACAAAGTTCATCTCCCT 297
 |||||
 328 CGGCTCTCTCTCCCTGTGGTGAACAAAGTGAAGATGCCAAGTCCATTAATTAAC 387
 |||||
 298 CCGCTCTCTCTCCCTGTGGTGAACAAAGTGAAGATGCCAAGTCCATTAATTAAC 357
 |||||
 388 GATCTTTATATGCAATCTCCAACTGGGGCCAGAGAGAAATGATGATGATGATGATG 447
 |||||
 358 GATCTTTATATGCAATCTCCAACTGGGGCCAGAGAGAAATGATGATGATGATGATG 417
 |||||
 448 ACAGACCCCACTCCCACTGT 507
 |||||
 418 ACTGACACCACTCCCACTGT 477
 |||||
 508 GAAAGAGAAAGATGAGAGATGAAGAGACCAAGCCCAAAATCATCTCAGACAGGAGA 567
 |||||
 478 GAAAGAGAA--ATGGAAGAAATGAGAGACTTAAGCCAAAATTTATCTCAGACAGGAGG 534

QY 568 CCGAGTACACACCGATCCACCTAGCTGA 597
 |||||
 DB 535 CCGAGTACACCGCGATCCACCTAGCTGA 564

RESULT 2
 US-08-081-610-3
 Sequence 3, Application US/08081610
 Patent No. 5445941
 GENERAL INFORMATION:
 APPLICANT: Yang, Na N
 TITLE OF INVENTION: Materials and Methods for Screening
 TITLE OF INVENTION: Anti-Osteoporosis or Serum Lipid Lowering Agents
 NUMBER OF SEQUENCES: 9
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Allergretti and Witcoff, Ltd.
 STREET: 10 S. Wacker Dr.
 CITY: Chicago
 STATE: IL
 COUNTRY: U.S.A
 ZIP: 60606
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/081,610
 FILING DATE:
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Heaphy, Barbara A
 REGISTRATION NUMBER: 34,619
 REFERENCE/DOCKET NUMBER: 93,402
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312-715-1000
 TELEFAX: 312-715-1234
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3303 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: mRNA
 LOCATION: 2170..3303
 FEATURE:
 NAME/KEY: mRNA
 LOCATION: 2214..3303
 FEATURE:
 NAME/KEY: mRNA
 LOCATION: 2219..3303
 FEATURE:
 NAME/KEY: misc RNA
 LOCATION: 3301..3303
 OTHER INFORMATION: /note="CDS Start, codon start = 1,
 OTHER INFORMATION: translation M"
 FEATURE:
 NAME/KEY: TATA_signal
 LOCATION: 2170..2176
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: 1896..2306
 OTHER INFORMATION: /note="pb-301 -301 to +110"
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: 1976..2306
 OTHER INFORMATION: /note="pb-221 -221 to +110"
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: 2106..2306
 OTHER INFORMATION: /note="pb-91 -91 to +110"

```

FEATURE:
NAME/KEY: misc feature
LOCATION: 2137..2306
OTHER INFORMATION: /note= "pB-60 -60 to +110"
FEATURE:
NAME/KEY: misc feature
LOCATION: 2150..2306
OTHER INFORMATION: /note= "pB-47 -47 to +110"
FEATURE:
NAME/KEY: misc feature
LOCATION: 2159..2306
OTHER INFORMATION: /note= "pB-38 -38 to +110"
FEATURE:
NAME/KEY: misc feature
LOCATION: 2159..2271
OTHER INFORMATION: /note= "TGFB-3 position -38 to
OTHER INFORMATION: +75"
FEATURE:
NAME/KEY: misc feature
LOCATION: 2159..2231
OTHER INFORMATION: /note= "TGFB-3 position -38 to
OTHER INFORMATION: +35"
US-08-081-610-3

```

```

Query Match      6.4%; Score 38.4; DB 1; Length 3303;
Best Local Similarity 58.9%; Pred. No. 0.015;
Matches 66; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

```

```

QY 421 GGAGAGAAGTGAATCTGACGCGAGACAGACCCCACTCCCACTGGTGGTCCACGTG 480
|||||
DB 2161 CGAGGAGGTATAAAATTCAGCAGAGAGAAATAGAGAAACAGTGTGTGCGATGTG 2220
|||||

```

```

QY 481 TGTGAGGTGACCAAGAGAAATGAGGAGGAGAGAGAGAGAGATGGAGAGATGA 532
|||||
DB 2221 TGTGTGAGAGAGAGGAGGAGGAGCGAGAGGAGGAGGAGGAGGAGAGA 2272
|||||

```

```

RESULT 3
US-08-232-463-14/c
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:

```

```

TELEPHONE: (703)836-9300
TELEFAX: (703)883-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PTZSPT-Fls
US-08-232-463-14

```

```

Query Match      6.3%; Score 37.8; DB 1; Length 7218;
Best Local Similarity 1.7%; Pred. No. 0.038;
Matches 3; Conservative 116; Mismatches 58; Indels 0; Gaps 0;

```

```

QY 409 AAGCTGGGCCAGAGAGAGAGTATGAATGTCATGACGAGCAGACAGACCCCACTCCCACTGGT 468
|||||
DB 1240 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1181
|||||

```

```

QY 469 GTGCTCCACGTGTGTGAGAGTACCAAGAGATGAGGAGAGAGAGAGATGGAGAGA 528
|||||
DB 1180 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1121
|||||

```

```

QY 529 ATGAAGAGACCCCAAGCCCAAAATCATCCAGACACGAGACCGGAGTACACCGATC 585
|||||
DB 1120 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1064
|||||

```

```

RESULT 4
US-09-056-105-31/c
Sequence 31, Application US/09056105
Patent No. 6287569
GENERAL INFORMATION:
APPLICANT: KIPPS, THOMAS J.
APPLICANT: WU, YUNQI
TITLE OF INVENTION: VACCINES WITH ENHANCED INTRACELLULAR
TITLE OF INVENTION: PROCESSING
FILE REFERENCE: 233/221
CURRENT APPLICATION NUMBER: US/09/056,105
CURRENT FILING DATE: 1998-04-06
EARLIER APPLICATION NUMBER: 60/043,467
EARLIER FILING DATE: 1997-04-10
NUMBER OF SEQ ID NOS: 35
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 31
LENGTH: 7130
TYPE: DNA
ORGANISM: Homo sapiens
US-09-056-105-31

```

```

Query Match      6.3%; Score 37.6; DB 4; Length 7130;
Best Local Similarity 61.0%; Pred. No. 0.044;
Matches 61; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

```

```

QY 483 TCAGAGTGCACCAAGAGATGAGGAGGAGAGAGATGAGAGATGAGAGAGATGAGAGAGCCAA 542
|||||
DB 3553 TGAGAGAAACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3494
|||||

```

```

QY 543 GCCCAAAATCATCCAGACACGAGACCGGAGTACACACCG 582
|||||
DB 3493 GGGAGAACAGTGACAAACAGGTGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3454
|||||

```

```

RESULT 5
US-08-913-832A-1
Sequence 1, Application US/08913832A
Patent No. 6329517
GENERAL INFORMATION:
APPLICANT: Seelig, Hans Peter
APPLICANT: Renz, Manfred
TITLE OF INVENTION: DERMATOMYOSITIS-SPECIFIC AUTO-ANTIGEN

```


TELEPHONE: (415)343-4341
TELEFAX: (415)343-4342
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2277 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-676-967-2

Query Match 5.6%; Score 33.2; DB 1; Length 2277;
Best Local Similarity 35.9%; Pred. No. 0.65;
Matches 51; Conservative 27; Mismatches 64; Indels 0; Gaps 0;

QY 422 GAGAGATGACTGCATGCGACGACACCCACCTCCAGTGTGTGTGTCACCTGT 481
DB 668 GNGAGARGARGAYATGARGARGARAAAGAYGAYGAYGAYGAYGAYGARG 727
QY 482 GTGAGAGTGACCAAGAGATGAGGAGAGAGAGATGAGAGATGAAGAGACCCA 541
DB 728 ARGAYGNGTNTTGGAYGARGARGARGARGARABAAATTHGARMSNAARGTNACNA 787
QY 542 AGCCCAAAATCATCCAGACACG 563
DB 788 ARCCNGTNCARATHCARAARMG 809

RESULT 11

US-08-676-974-2
Sequence 2, Application US/08676974
Patent No. 5770422

GENERAL INFORMATION:
APPLICANT: COLLINS, KATHLEEN
TITLE OF INVENTION: Human Telomerase
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Science & Technology Law Group
STREET: 268 Bush Street, Suite 3200
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/676,974
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Osman Ph.D., Richard A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UCB96-055
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)343-4341
TELEFAX: (415)343-4342
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2277 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-676-974-2

Query Match 5.6%; Score 33.2; DB 1; Length 2277;
Best Local Similarity 35.9%; Pred. No. 0.65;
Matches 51; Conservative 27; Mismatches 64; Indels 0; Gaps 0;
QY 422 GAGAGATGACTGCATGCGACGACACCCACCTCCAGTGTGTGTGTCACCTGT 481

DB 668 GNGAGARGARGAYATGARGARGARAAAGAYGAYGAYGAYGAYGARG 727
QY 482 GTGAGAGTGACCAAGAGATGAGGAGAGAGAGATGAGAGATGAAGAGACCCA 541
DB 728 ARGAYGNGTNTTGGAYGARGARGARGARGARABAAATTHGARMSNAARGTNACNA 787
QY 542 AGCCCAAAATCATCCAGACACG 563
DB 788 ARCCNGTNCARATHCARAARMG 809

RESULT 12
US-09-098-487-2
Sequence 2, Application US/09098487
Patent No. 5917025

GENERAL INFORMATION:
APPLICANT: COLLINS, Kathleen
TITLE OF INVENTION: Human Telomerase
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Science & Technology Law Group
STREET: 268 Bush Street, Suite 3200
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/098,487
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Osman Ph.D., Richard A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UCB96-055
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)343-4341
TELEFAX: (415)343-4342
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2277 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-098-487-2

Query Match 5.6%; Score 33.2; DB 2; Length 2277;
Best Local Similarity 35.9%; Pred. No. 0.65;
Matches 51; Conservative 27; Mismatches 64; Indels 0; Gaps 0;

QY 422 GAGAGATGACTGCATGCGACGACACCCACCTCCAGTGTGTGTGTCACCTGT 481
DB 668 GNGAGARGARGAYATGARGARGARAAAGAYGAYGAYGAYGAYGARG 727
QY 482 GTGAGAGTGACCAAGAGATGAGGAGAGAGAGATGAGAGATGAAGAGACCCA 541
DB 728 ARGAYGNGTNTTGGAYGARGARGARGARGARABAAATTHGARMSNAARGTNACNA 787
QY 542 AGCCCAAAATCATCCAGACACG 563
DB 788 ARCCNGTNCARATHCARAARMG 809

RESULT 13
US-09-312-038-3/c
Sequence 3, Application US/09312038A
Patent No. 6232526

FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR93/00623
FILING DATE: 23-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 92/07785
FILING DATE: 25-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Julie K.
REGISTRATION NUMBER: 38,619
REFERENCE/DOCKET NUMBER: ST92040-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610)454-3839
TELEFAX: (610)454-3808
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1685 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Kluyveromyces fragilis
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1683
OTHER INFORMATION: /product= "Protease B gene"
OTHER INFORMATION: /gene= "K1.PRB1"

US-08-360-673-1

Query Match 5.5%; Score 33; DB 1; Length 1685;
Best Local Similarity 53.5%; Pred. No. 0.64;
Matches 69; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

OY 9 TAGGACTTTAGCTACATTTTAACTCCCTGATGCTTGTGTGGCAACGATGATGCTT 68
DB 595 TAGAATCTTCTTCAACAAATTCACCCATGCGCTTCTTCTGATGAATGATTAATCTT 536
OY 69 CAGCGAAAGTGAAGCAGGCGCAATTTGATCCCTCTTCAGAACATATGACAGGACAC 128
DB 535 TTGTGAAGTAGCCACATTAACGAGAGGAAACCGCAATATCAAAAGAGCTTGAATAC 476
OY 129 CACCTTCCA 137
DB 475 CACCTTCCA 467

Search completed: December 14, 2002, 18:14:01
Job time : 39.45 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2002, 15:19:20 ; Search time 33.1761 Seconds
(without alignments)
13100.560 Million cell updates/sec

Title: US-09-782-953-5

Perfect score: 597

Sequence: 1 atggattttgggacttag.....caccgatccaccattagctga 597

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 422006 seqs, 364007625 residues

Total number of hits satisfying chosen parameters: 844012

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents_NA New:*

1: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:*

2: /cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq:*

3: /cgn2_6/ptodata/1/pna/US07_NEW_COMB.seq:*

4: /cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq:*

5: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:*

6: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:*

7: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	470.6	78.8	1931	5	US-09-724-676-12748 Sequence 12748, A
2	470.6	78.8	1931	5	US-09-724-676A-12748 Sequence 12748, A
3	470.6	78.8	2050	5	US-09-724-676-12756 Sequence 12756, A
4	470.6	78.8	2050	5	US-09-724-676A-12756 Sequence 12756, A
5	470.6	78.8	2297	5	US-09-724-676-12747 Sequence 12747, A
6	470.6	78.8	2297	5	US-09-724-676A-12747 Sequence 12747, A
7	459.6	77.0	2348	6	US-10-240-965-60 Sequence 60, Appl
8	408.2	68.4	1893	5	US-09-724-676-12741 Sequence 12741, A
9	408.2	68.4	1893	5	US-09-724-676A-12741 Sequence 12741, A
10	401.6	67.3	1839	5	US-09-724-676-12743 Sequence 12743, A
11	401.6	67.3	1839	5	US-09-724-676A-12743 Sequence 12743, A
12	401.6	67.3	1876	5	US-09-724-676-12745 Sequence 12745, A
13	401.6	67.3	1876	5	US-09-724-676A-12745 Sequence 12745, A
14	350.6	58.7	911	5	US-09-724-676-12749 Sequence 12749, A
15	350.6	58.7	911	5	US-09-724-676A-12749 Sequence 12749, A
16	348.2	58.3	1094	5	US-09-724-676-12750 Sequence 12750, A
17	348.2	58.3	1094	5	US-09-724-676A-12750 Sequence 12750, A
18	348.2	58.3	1213	5	US-09-724-676-12740 Sequence 12740, A
19	348.2	58.3	1213	5	US-09-724-676A-12740 Sequence 12740, A
20	285.8	47.9	1056	5	US-09-724-676-12742 Sequence 12742, A
21	285.8	47.9	1056	5	US-09-724-676A-12742 Sequence 12742, A
22	279.2	46.8	1002	5	US-09-724-676-12744 Sequence 12744, A
23	279.2	46.8	1002	5	US-09-724-676A-12744 Sequence 12744, A
24	279.2	46.8	1039	5	US-09-724-676-12746 Sequence 12746, A
25	279.2	46.8	1039	5	US-09-724-676A-12746 Sequence 12746, A
26	273.8	45.9	442	5	US-09-513-999C-1772 Sequence 1772, Ap

27	215.8	36.1	2411	5	US-09-724-676-15329	Sequence 15329, A
28	215.8	36.1	2411	5	US-09-724-676A-15329	Sequence 15329, A
29	209.6	35.1	3159	6	US-10-290-438-1	Sequence 1, Appli
30	191	32.0	1577	5	US-09-724-676-12755	Sequence 12755, A
31	191	32.0	1577	5	US-09-724-676A-12755	Sequence 12755, A
32	191	32.0	1943	5	US-09-724-676-12752	Sequence 12752, A
33	191	32.0	1943	5	US-09-724-676A-12752	Sequence 12752, A
34	186.2	31.2	615	6	US-10-290-438-8	Sequence 8, Appli
35	155	26.0	1021	6	US-10-290-438-3	Sequence 3, Appli
36	144.6	24.2	446	6	US-10-203-138A-2101	Sequence 2101, Ap
37	123	20.6	486	6	US-10-203-138A-827	Sequence 827, App
38	84.4	14.1	123	6	US-10-203-138A-5980	Sequence 5980, Ap
39	72	12.1	305	5	US-09-724-676-12751	Sequence 12751, A
40	72	12.1	305	5	US-09-724-676A-12751	Sequence 12751, A
41	71	11.9	557	5	US-09-724-676-12753	Sequence 12753, A
42	71	11.9	557	5	US-09-724-676A-12753	Sequence 12753, A
43	70.6	11.8	85	6	US-10-203-138A-7216	Sequence 7216, Ap
44	68.6	11.5	740	5	US-09-724-676-12754	Sequence 12754, A
45	68.6	11.5	740	5	US-09-724-676A-12754	Sequence 12754, A

ALIGNMENTS

RESULT 1

US-09-724-676-12748
; Sequence 12748, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12748
; LENGTH: 1931
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (11)..(11)
; OTHER INFORMATION: n is a,c,g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (18)..(18)
; OTHER INFORMATION: n is a,c,g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (23)..(23)
; OTHER INFORMATION: n is a,c,g, or t
US-09-724-676-12748

Query Match 78.8%; Score 470.6; DB 5; Length 1931;
Best Local Similarity 87.9%; Pred. No. 3e+128;
Matches 525; Conservative 0; Mismatches 69; Indels 3; Gaps 1;

Qy	1	ATGGATTTTGGGACTTTAGCTACAAATTTAGCTCCCTGATTGCTTGTGCGCAACGAT	60
Db	63	ATGCATTTTGGGAACTTTAACTACAGTTTGTAGTCCCTGATTGCTGCGCAACAGT	122
Qy	61	GATGTCCTTACGCGAAGTGAGACCCAGGCGCAAAATTTGAATCCCTCTTCAGAACATATGAC	120
Db	123	GATATCTTTCAGCGAAGTGAAACCCAGGCGCAAAATTTGAGTCCCTCTTTAGGACGATGAC	182
Qy	121	AAGGACACACTTCCAGTATTTAAAGAGCTTCAAAAGTGTCCGGAATAAATTCAGCAAC	180
Db	183	AAGGACATCACCTTTCAGTATTTTAAAGAGCTTCAAAAGAGTCAAGATAAATTCAGCAAC	242
Qy	181	CCCTTATCTGAGCGGAGTCCGAGCTCGCGCTGCACAGACCGAGTTCCTGGGGAAGGA	240
Db	243	CCCTTCTCCGAGCAGATGCCAGGCTCCAGCTGCATAGACTGAGTTCTTGGGGAAGGA	302

QY 241 ATGAGTTGATTTTGTCTGACACTTACACATAGAGAGTTGACACTGCTCCGCCCAAT 300
 DB 303 ATGAGTTATATTTTGTCTGACACTTACACATAGAGAGTTGACACTGCTCCGCCCAAT 362
 QY 301 CCCGACAAACAGTTCTCTGATCTCCCTCCGCGCTCTCTCCCTGCTGGAGAACAGTA 360
 DB 363 CCAGACAAAGCAGTTCTGATCTCCCTCCGCGCTCTCCGCGAGTGGAGAACAGTGTG 422
 QY 361 GAAGATGCAACCCCGCTGATCTCTGATCTCTGATCTCTGATCTCTGATCTCTGATCT 420
 DB 423 GAAGATGCAACCCCGCTGATCTCTGATCTCTGATCTCTGATCTCTGATCTCTGATCT 482
 QY 421 GGAGAGAGTATGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 480
 DB 483 GGGGAAAGTATGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 542
 QY 481 TGTGAGAGTACCAAGAGATGAGAGAGAGAGAGAGAGATGAGAGAGATGAGAGAGAGCC 540
 DB 543 TGTGAGAGTACCAAGAGATGAGAGAGAGAGAGAGAGATGAGAGAGATGAGAGAGAGCT 599
 QY 541 AAGCCCAAAATCATCTCAGACACGAGAGACCGAGTACACACCTTCACTTACCTGA 597
 DB 600 AAGCCCAAAATTTATCTCAGACACGAGAGACCGAGTACACACCTTCACTTACCTGA 656

RESULT 2

US-09-724-676A-12748
 ; Sequence 12748, Application US/09724676A
 ; GENERAL INFORMATION:
 ; APPLICANT: Compugen LTD
 ; TITLE OF INVENTION: Variants of alternative splicing
 ; FILE REFERENCE: 129181.4 Compugen
 ; CURRENT APPLICATION NUMBER: US/09/724,676A
 ; CURRENT FILING DATE: 2000-11-28
 ; NUMBER OF SEQ ID NOS: 97222
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 12748
 ; LENGTH: 1931
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (11)..(11)
 ; OTHER INFORMATION: n is a,c,g, or t
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (18)..(18)
 ; OTHER INFORMATION: n is a,c,g, or t
 ; NAME/KEY: misc_feature
 ; LOCATION: (23)..(23)
 ; OTHER INFORMATION: n is a,c,g, or t
 ; US-09-724-676A-12748

Query Match 78.8%; Score 470.6; DB 5; Length 1931;

Best Local Similarity 87.9%; Pred. No. 3e-128; Mismatches 69; Indels 3; Gaps 1;

QY 1 ATGAGTTTGGGACCTTAGTACATTTTAGCTCCCTGATGCTGCTGAGGCAAAAGAT 60
 DB 63 ATGAGTTTGGGACCTTAGTACATTTTAGCTCCCTGATGCTGCTGAGGCAAAAGAT 122
 QY 61 GATGCTTCAGCGAAAGTGAAGCCAGGCGCAAAATTTGAATCCCTCTTCAAGACATATGAC 120
 DB 123 GATGCTTCAGCGAAAGTGAAGCCAGGCGCAAAATTTGAATCCCTCTTCAAGACATATGAC 182
 QY 121 AAGGACACCACTTCAGATTTTAAAGCTTCAAGCTGTGCGGATTAATCTCAGAAC 180
 DB 183 AAGGACATCACTTCAGATTTTAAAGCTTCAAGCTGTGCGGATTAATCTCAGAAC 242
 QY 181 CCCTTATCTGACCGAGTGCAGGCTGCGGCTGCACAAGACCGAGTTCCTGGGAGAGAA 240
 DB 243 CCCTTCTCGAGCAGATGCCAGGCTCCAGCTGCATTAAGACTGAGTTTCTGGGAAAGAA 302

QY 241 ATGAGTTGATTTTGTCTGACACTTACACATAGAGAGTTGACACTGCTCCGCCCAAT 300
 DB 303 ATGAGTTATATTTTGTCTGACACTTACACATAGAGAGTTGACACTGCTCCGCCCAAT 362
 QY 301 CCCGACAAACAGTTCTCTGATCTCCCTCCGCGCTCTCTCCCTGCTGGAGAACAGTA 360
 DB 363 CCAGACAAAGCAGTTCTGATCTCCCTCCGCGCTCTCTCCGCGAGTGGAGAACAGTGTG 422
 QY 361 GAAGATGCAACCCCGCTGATCTCTGATCTCTGATCTCTGATCTCTGATCTCTGATCT 420
 DB 423 GAAGATGCAACCCCGCTGATCTCTGATCTCTGATCTCTGATCTCTGATCTCTGATCT 482
 QY 421 GGAGAGAGTATGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 480
 DB 483 GGGGAAAGTATGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 542
 QY 481 TGTGAGAGTACCAAGAGATGAGAGAGAGAGAGAGAGATGAGAGAGATGAGAGAGAGCC 540
 DB 543 TGTGAGAGTACCAAGAGATGAGAGAGAGAGAGAGAGATGAGAGAGATGAGAGAGAGCT 599
 QY 541 AAGCCCAAAATCATCTCAGACACGAGAGACCGAGTACACACCTTCACTTACCTGA 597
 DB 600 AAGCCCAAAATTTATCTCAGACACGAGAGACCGAGTACACACCTTCACTTACCTGA 656

RESULT 3

US-09-724-676-12756
 ; Sequence 12756, Application US/09724676
 ; GENERAL INFORMATION:
 ; APPLICANT: Compugen LTD
 ; TITLE OF INVENTION: Variants of alternative splicing
 ; FILE REFERENCE: 129181.4 Compugen
 ; CURRENT APPLICATION NUMBER: US/09/724,676
 ; CURRENT FILING DATE: 2000-11-28
 ; NUMBER OF SEQ ID NOS: 97222
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 12756
 ; LENGTH: 2050
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (7)..(7)
 ; OTHER INFORMATION: n is a,c,g, or t
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (47)..(47)
 ; OTHER INFORMATION: n is a,c,g, or t
 ; US-09-724-676-12756

Query Match 78.8%; Score 470.6; DB 5; Length 2050;

Best Local Similarity 87.9%; Pred. No. 3.1e-128; Mismatches 69; Indels 3; Gaps 1;

QY 1 ATGAGTTTGGGACCTTAGTACATTTTAGCTCCCTGATGCTGCTGAGGCAAAAGAT 60
 DB 182 ATGAGTTTGGGACCTTAGTACATTTTAGCTCCCTGATGCTGCTGAGGCAAAAGAT 241
 QY 61 GATGCTTCAGCGAAAGTGAAGCCAGGCGCAAAATTTGAATCCCTCTTCAAGACATATGAC 120
 DB 242 GATGCTTCAGCGAAAGTGAAGCCAGGCGCAAAATTTGAATCCCTCTTCAAGACATATGAC 301
 QY 121 AAGGACACCACTTCAGATTTTAAAGCTTCAAGCTGTGCGGATTAATCTCAGAAC 180
 DB 302 AAGGACATCACTTCAGATTTTAAAGCTTCAAGCTGTGCGGATTAATCTCAGAAC 361
 QY 181 CCCTTATCTGACCGAGTGCAGGCTGCGGCTGCACAAGACCGAGTTCCTGGGAGAGAA 240
 DB 362 CCCTTCTCGAGCAGATGCCAGGCTCCAGCTGCATTAAGACTGAGTTTCTGGGAAAGAA 421
 QY 241 ATGAGTTGATTTTGTCTGACACTTACACATAGAGAGTTGACACTGCTCCGCCCAAT 300

Db 422 ATGAAGTTATATTTTGTCTCAGACCTTACATAGGAAGCTCAGACCTGGCTCCGCCAAAT 481
Qy 301 CCCGACAAACAGTTCTCTCATCTCCCTCCGCGCTCTCTCCGTTGGCTGGAACAAGTA 360
Db 482 CCAGACAAAGAGTTTCTGATCTCCCTCCGCGCTCTCCGCGAGTGGATGGAACAAGTG 541
Qy 361 GAAGATGCCACCCCGCTCATAAATTTAGCATCTTTATATGCACTCTCAAGCTGGGGCCA 420
Db 542 GAAGTGGACCCAGTCATAACTATGATCTCTTATATGCCATCTCAAGCTGGGGCCA 601
Qy 421 GGAGAGAATGATGAATGTCATGCGACGACAGACCCCACTCCAGTGTGGTCCACGTG 480
Db 602 GGGGAAAAGTATGAATTCACGCGAGCTGACACCACTCCACGCTGGTCCATGTA 661
Qy 481 TGTGAGAGTACCAAGAGATCAGGAGGAGGAGAGAGATGGAGAGATGAAGAGACCC 540
Db 662 TGTGAGAGTATCAAGAGAGGAGGAGGAGGAGGAA---ATGGAAGAAATGAGGAGACCT 718
Qy 541 AAGCCCAAAATCATCCAGACACGAGACCGGAGTACACACCGATCCACCTTAGCTGA 597
Db 719 AAGCCAAAATTTATCCAGACAGGAGGCGGAGTACACCGGATCCACCTCAGCTGA 775

RESULT 4

US-09-724-676A-12756
; Sequence 12756, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 12756
; LENGTH: 2050
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (7)..(7)
; FEATURE:
; OTHER INFORMATION: n is a,c,g, or t
; NAME/KEY: misc feature
; LOCATION: (47)..(47)
; OTHER INFORMATION: n is a,c,g, or t
US-09-724-676A-12756

Query Match 78.8%; Score 470.6; DB 5; Length 2050;
Best Local Similarity 87.9%; Pred. No. 3.1e-128;
Matches 525; Conservative 0; Mismatches 69; Indels 3; Gaps 1;

Qy 1 ATGGAATTTAGGAGCTTTAGCTACAAATTTTAGCTCCCTGATTCCTTGTGTGGCAACAGAT 60
Db 182 ATGCAATTTAGAACTTTAACTACAGTTTTAGCTCCCTGATTCCTTGTGTGGCAACAGAT 241
Qy 61 GATGCTTTCAGCGAAGTGAACAGGCGGCAAAATTTGAATCCCTCTTCAGAACATATGAC 120
Db 242 GATATCTTCAGCGAAGTGAACAGGCGGCAAAATTTGAGTCCCTCTTTAGGAGCTATGAC 301
Qy 121 AAGGACACCACTTCCAGTATTTTAAGAGCTTCAAAAGCTGTCGGGATAAACTTCAGCAAC 180
Db 302 AAGGACATCACTTTTCAGTATTTTAAGAGCTTCAAAAGCTGTCAGAACTTCAGCAAC 361
Qy 181 CCCTTATCTCAGCGAGTCCAGGCTCGCGCTGCAAGACCGAGTTCCTGGGGAGGAA 240
Db 362 CCCTTCTCCGAGCAGATGCCAGGCTCCAGCTGCATAAGACTGAGTTTCTGGGAAGGAA 421
Qy 241 ATGAAGTTGATTTTCTCAGACTTTTACACATAGGAAGTTTCAACCTGGCTCGGCCAAAT 300
Db 422 ATGAAGTTATATTTTGTCTCAGACTTTTACACATAGGAAGCTTCACCTGGCTCGGCCAAAT 481
Qy 301 CCCGACAAACAGTTTCTCTCATCTCCCTCCGCGCTCTCTCCCTGGCTGGGTAACAAGTA 360

Db 482 CCAGACAAAGCAGTTTCTGATCTCCCTCCGCGCTCTCCGCGAGTGGATGGAACAAGTG 541
Qy 361 GAAGATGCCACCCCGCTCATAAATTTAGCATCTTTATATGCACTCTCAAGCTGGGGCCA 420
Db 542 GAAGATGCCACCCCGCTCATAACTATGATCTCTTATATGCCATCTCCAAAGCTGGGGCCA 601
Qy 421 GGAGAGAATGATGAATGTCATGCGACGACAGACCCCACTCCAGTGTGGTCCACGTG 480
Db 602 GGGGAAAAGTATGAATTCACGCGAGCTGACACCACTCCACGCTGGTCCATGTA 661
Qy 481 TGTGAGAGTACCAAGAGATGAGGAGGAGGAGAGATGGAGAGATGAAGAGACCC 540
Db 662 TGTGAGAGTATCAAGAGAGGAGGAGGAGGAGGAA---ATGGAAGAAATGAGGAGACCT 718
Qy 541 AAGCCCAAAATCATCCAGACACGAGACCGGAGTACACACCGATCCACCTTAGCTGA 597
Db 719 AAGCCAAAATTTATCCAGACAGGAGGCGGAGTACACCGGATCCACCTCAGCTGA 775

RESULT 5

US-09-724-676-12747
; Sequence 12747, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 12747
; LENGTH: 2297
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (11)..(11)
; OTHER INFORMATION: n is a,c,g, or t
; NAME/KEY: misc feature
; LOCATION: (18)..(18)
; OTHER INFORMATION: n is a,c,g, or t
; NAME/KEY: misc feature
; LOCATION: (23)..(23)
; OTHER INFORMATION: n is a,c,g, or t
US-09-724-676-12747

Query Match 78.8%; Score 470.6; DB 5; Length 2297;
Best Local Similarity 87.9%; Pred. No. 3.2e-128;
Matches 525; Conservative 0; Mismatches 69; Indels 3; Gaps 1;

Qy 1 ATGGAATTTAGGAGCTTTAGCTACAAATTTTAGCTCCCTGATTCCTTGTGTGGCAACAGAT 60
Db 63 ATGCAATTTAGAACTTTAACTACAGTTTTAGCTCCCTGATTCCTTGTGTGGCAACAGAT 122
Qy 61 GATGCTTTCAGCGAAGTGAACAGGCGGCAAAATTTGAATCCCTCTTCAGAACATATGAC 120
Db 123 GATATCTTCAGCGAAGTGAACAGGCGGCAAAATTTGAGTCCCTCTTTAGGAGCTATGAC 182
Qy 121 AAGGACACCACTTCCAGTATTTTAAGAGCTTCAAAAGCTGTCGGGATAAACTTCAGCAAC 180
Db 183 AAGGACATCACTTTTCAGTATTTTAAGAGCTTCAAAAGCTGTCAGAACTTCAGCAAC 242
Qy 181 CCCTTATCTCAGCGAGTCCAGGCTCGCGCTGCAAGACCGAGTTCCTGGGGAGGAA 240
Db 243 CCCTTCTCCGAGCAGATGCCAGGCTCCAGCTGCATAAGACTGAGTTTCTGGGAAGGAA 302
Qy 241 ATGAAGTTGATTTTGTCTCAGACTTTTACACATAGGAAGTTTCAACCTGGCTCGGCCAAAT 300
Db 303 ATGAAGTTATATTTTGTCTCAGACTTTTACACATAGGAAGCTTCACACCTGGCTCGGCCAAAT 362

Query Match	Similarity	78.8%	Score 470.6	DB 5	Length 2297
Best Local	Similarity	87.9%	Pred. No. 3,2e-128		
Matches 525	Conservative	0	Mismatches 69	Indels 3	Gaps 1
Qy	1	ATGATATTTAGGACTTTAGCTACAAATTTAGCTCCCTGATGCTGTTGGCAAAAGAT	60		
Db	63	ATCCATTTTAGAATCTTAAGTTTAACTGATTTAGCTCCCTGATGCTGTTGGCAAAAGAT	122		
Qy	61	GATGCTTCAGCGAAAGTAGACCAAGGCCAAATTTGAATCCCTCTTCAGAAATATGAC	120		
Db	123	GATATCTTCAGCAAAAGTAGAAACCAAGGCCAAATTTGATGCTCCCTCTTTAGAGATATGAC	182		
Qy	121	AAGGACACCACTTCCTCCGATATTTTAAGAGCTTCAAACGTTCCGGAAATTAACCTCAGCAAC	180		
Db	183	AAGGACATCACTCTTCATGATTTTAAAGCTTCAAAGAGCTCAGAAATTAACCTCAGCAAC	242		
Qy	181	CCCTTAATGACAGCGATGACAGGCTCGGCTGCAACAGCCAGTCTCTGGGGAAGAA	240		
Db	243	CCCTTCTCCGAGAGATGCAAGCTTCAGCTGATTAAGCTAGTTTCTGGGGAAGAA	302		
Qy	241	ATGAGTTGATTTTGTCTCAGACTTTACATAGGAAGTTACACCTCGGCTCCGCCAAT	300		
Db	303	ATGAGTTTATTTTGTCTCAGACTTTACATAGGAAGCTACACCTCGGCTCCGCCAAT	362		

Query Match:	77.0%	Score 459.6;	DB 6;	Length 2348;
Best Local Similarity:	87.8%	Pred. No. 5,5e-125;		
Matches 555;	Conservative 0;	Mismatches 63;	Indels 4;	Gaps 2;

Qy	1	ATGGAATTTAGGACCTTTAGTACAATTTTAACTCCCTGATGTCCTTGTTGTGCGAAACGAT	60
Db	151	ATGCATTTTAGAACTTTAACTACAGATTTTAACTCCCTGATGTCCTGTTGTGCGAAACACT	210
Qy	61	GATGCTTTCAGCGCAAGTGAACCAAGGCGCAATTGAAATCCCTCTCGAACATATGAC	120
Db	211	GATATCTTCAGCGCAAGTGAACCAAGGCGCAATTGAACTCCCTCTTGAACGATGAC	270
Qy	121	AAGGACACCACCTTCAGTATTTTAAAGACTTCAAACTGTCCGGATAAATCTTCAGCAAC	180
Db	271	AAGGACATCACCTTCAAGTATTTTAAAGACTTCAAAACGATCAGAAATTAACCTTCAGCAAC	330
Qy	181	CCCTTAATTCGACCGCAGTCGCGGCTGCGACACAGACCGAGTTCCTGGGGAAGAA	240
Db	331	CCCTTCCTCGCAGACAGATGCGAGGCTCCAGCTGCAATMAACATGAGTTTCTGGGAAAGGAA	390
Qy	241	ATGAAGTTGATTTTCTCAGACTTACACATGAGAAATTTCACCTT-GAGTCCGCGCAA	299

Db 391 ATGAAGTTATATTTTGTCTCAGACCTTACATAGGAAGCTCACACCTGGGCTCCGCCAAA 450
Qy 300 TCCCGCAAAACAGTTCTCATCTCCCTCCGCGCTCTCTCCCGTTGGCTGGAACAAAGT 359
Db 451 TCCGACAAAGCAGTTTCTGATCTCCCTCCGCGCTCTCCGCGAGTGGATGGAACAAAGT 510
Qy 360 AGAAGATCCACCCCGCTCATAAATTAGATCTTTTATATGCCATCTCCAAAGCTGGGGCC 419
Db 511 GGAAGATCCGACCCCGCTCATAAATTAGATCTTTTATATGCCATCTCCAAAGCTGGGGCC 570
Qy 420 AGGAGAGATGATGAATGCAATGCGAGCGACAGACCCCACTCCAGTGGTGGTCCAGGT 479
Db 571 AGGGGAAAGATGATGATGCGAGCGAGCGACTGACACCACTCCAGCGTGGTGGTCCATGT 630
Qy 480 GTGTGAGAGTGAACCAAGAGATGAGGAGGAGAGAGAGAGAGATGGAGAGATGAAGAGACC 539
Db 631 ATGTGAGAGTGAATCAAGAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 687
Qy 540 CAAGCCCAAAATCACCAGACCGGAGACCGGAGTACACCGATCCACCTTACCTTACCTGA 597
Db 688 TAAGCCAAAATTTATCCAGACCGAGGCGCGAGTACACCGGATCCACCTCAGCTGA 745

RESULT 8
US-09-724-676-12741
; Sequence 12741, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 12741
; LENGTH: 1893
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676-12741

Query Match 68.4%; Score 408.2; DB 5; Length 1893;
Best Local Similarity 81.4%; Pred. No. 6.1e-110;
Matches 486; Conservative 0; Mismatches 108; Indels 3; Gaps 1;

Qy 1 ATGAATTTAGGAGCTTAGCTACAAATTTTGTCTCCCTGATGCTTGTGTGGCAACGAT 60
Db 25 ATGGAGGAGTGGACCTCGAGGACCTCCCGAGCGCCACCATCGCTGTACCTGGACCCG 84
Qy 61 GATGCTTTCAGCGAAGTGAAG 120
Db 85 CGCGTGTTCGTGGACGGCTGTGCGGGCCAAATTTGAGTCCCTCTTTAGGACGTATGAC 144
Qy 121 AAGGACACCACTTCCAGTATTTTAAAGAGCTTCAAAAGCTGTCGGATAACTTTCAGCAAC 180
Db 145 AAGGACATCACTTTCAGTATTTTAAAGAGCTTCAAAAGCTGAGATTAATCTTCAGCAAC 204
Qy 181 CCCTTATCTGAGCGAGTCCAGGCTCGGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
Db 205 CCCTTCTCGCAGCAGATGCCAGGCTCCAGCTGCATAGAGCTGAGTTTCTGGGAAAGGAA 264
Qy 241 ATGAAGTTGATTTTGTCTCAGACTTTTACATAGGAAGTTTACACCTGGCTCCGCCCAAT 300
Db 265 ATGAAGTTATATTTTGTCTCAGACTTTTACATAGGAAGTTTACACCTGGCTCCGCCCAAT 324
Qy 301 CCCGACAAACAGTTCTCATCTCCCTCCGCGCTCTCCCTCCGCTTGGCTGGGAAACAAGTA 360
Db 325 CCAGACAAAGCAGTTTCTGATCTCCCTCCGCGCTCTCCGCGAGTGGATGGAACAAGTG 384
Qy 361 GAAGATGCGACCCCGCTCATAAATTTACGATCTTTTATGCGATCTCCAGCTGGGGCA 420
Db 385 GAAGATGCGACCCCGCTCATAAATTTACGATCTTTTATGCGATCTCCAGCTGGGGCA 444
Qy 421 ATGAAGTTGATTTTGTCTCAGACTTTTACATAGGAAGTTTACACCTGGCTCCGCCCAAT 300
Db 265 ATGAAGTTATATTTTGTCTCAGACTTTTACATAGGAAGTTTACACCTGGCTCCGCCCAAT 324
Qy 301 CCCGACAAACAGTTCTCATCTCCCTCCGCGCTCTCCCTCCGCTTGGCTGGGAAACAAGTA 360
Db 325 CCAGACAAAGCAGTTTCTGATCTCCCTCCGCGCTCTCCGCGAGTGGATGGAACAAGTG 384
Qy 361 GAAGATGCGACCCCGCTCATAAATTTACGATCTTTTATGCGATCTCCAGCTGGGGCA 420
Db 385 GAAGATGCGACCCCGCTCATAAATTTACGATCTTTTATGCGATCTCCAGCTGGGGCA 444
Qy 421 GGAGAGAGATGATGAATGCAATGCGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480

Db 445 GGGGAAAAGTATGAATTTGACGAGCGACTGACACCACTCCCGAGCTGGTGGTCCATGTA 504
Qy 481 TGTGAGAGTGACCAAGAGATGAGGAGGAGAGAGATGGAGAGATGAAGAGAGACC 540
Db 505 TGTGAGAGTGATCAAGAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 561
Qy 541 AAGCCCAAAATCACCAGACCGGAGTACACACCGATCCACCTTACCTTACCTGA 597
Db 562 AAGCCCAAAATTTATCCAGACCGAGGCGCGAGTACACCGCGATCCACCTCAGCTGA 618

RESULT 9
US-09-724-676A-12741
; Sequence 12741, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 12741
; LENGTH: 1893
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676A-12741

Query Match 68.4%; Score 408.2; DB 5; Length 1893;
Best Local Similarity 81.4%; Pred. No. 6.1e-110;
Matches 486; Conservative 0; Mismatches 108; Indels 3; Gaps 1;

Qy 1 ATGAATTTAGGAGCTTAGCTACAAATTTTGTCTCCCTGATGCTTGTGTGGCAACGAT 60
Db 25 ATGGAGGAGTGGACCTCGAGGACCTCCCGAGCGCCACCATCGCTGTACCTGGACCCG 84
Qy 61 GATGCTTTCAGCGAAGTGAAG 120
Db 85 CGCGTGTTCGTGGACGGCTGTGCGGGCCAAATTTGAGTCCCTCTTTAGGACGTATGAC 144
Qy 121 AAGGACACCACTTCCAGTATTTTAAAGAGCTTCAAAAGCTGTCGGATAACTTTCAGCAAC 180
Db 145 AAGGACATCACTTTCAGTATTTTAAAGAGCTTCAAAAGCTGAGATTAATCTTCAGCAAC 204
Qy 181 CCCTTATCTGAGCGAGTCCAGGCTCGGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
Db 205 CCCTTCTCGCAGCAGATGCCAGGCTCCAGCTGCATAGAGCTGAGTTTCTGGGAAAGGAA 264
Qy 241 ATGAAGTTGATTTTGTCTCAGACTTTTACATAGGAAGTTTACACCTGGCTCCGCCCAAT 300
Db 265 ATGAAGTTATATTTTGTCTCAGACTTTTACATAGGAAGTTTACACCTGGCTCCGCCCAAT 324
Qy 301 CCCGACAAACAGTTCTCATCTCCCTCCGCGCTCTCCCTCCGCTTGGCTGGGAAACAAGTA 360
Db 325 CCAGACAAAGCAGTTTCTGATCTCCCTCCGCGCTCTCCGCGAGTGGATGGAACAAGTG 384
Qy 361 GAAGATGCGACCCCGCTCATAAATTTACGATCTTTTATGCGATCTCCAGCTGGGGCA 420
Db 385 GAAGATGCGACCCCGCTCATAAATTTACGATCTTTTATGCGATCTCCAGCTGGGGCA 444
Qy 421 GGAGAGAGATGATGAATGCAATGCGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
Db 445 GGGGAAAAGTATGAATTTGACGAGCGACTGACACCACTCCCGAGTGGTGGTCCATGTA 504
Qy 481 TGTGAGAGTGACCAAGAGATGAGGAGGAGAGAGATGGAGAGATGAAGAGAGACC 540
Db 505 TGTGAGAGTGATCAAGAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 561
Qy 541 AAGCCCAAAATCACCAGACCGGAGTACACACCGATCCACCTTACCTTACCTGA 597
Db 562 AAGCCCAAAATTTATCCAGACCGAGGCGCGAGTACACCGCGATCCACCTCAGCTGA 618

QY 206 TGGCGCTGCACAGACCGAGTTCCTGGGAGGAAATGAAGTTGATTTTGTCTCAGACTT 265
Db 213 TCCAGCTGCATAGACAGTGGTTCTGGGAAAGGAAATGAAGTTATATTTTGTCTCAGACTT 272
QY 266 TACACATAGGAAGTTACACCTGGCTCGGCCAATCCGCAACAGTTCTCTCATCTCCC 325
Db 273 TACACATAGGAAGTTACACCTGGCTCGGCCAATCCGCAACAGTTTCTGATCTCCC 332
QY 326 CTCCGGCTCTCTCCGTTGGCTGGAAACAAAGTAGAGATGCCACCCCGTGCATAAAT 385
Db 333 CTCCGGCTCTCTCCGTTGGCTGGAAACAAAGTAGAGATGCCACCCCGTGCATAAAT 392
QY 386 ACGATCTTTATATGTCATCTCCAGCTGGGGCCAGAGAGATGAATGCTGCAG 445
Db 393 ATGATCTCTTATATGTCATCTCCAGCTGGGGCCAGAGAGATGAATGCTGCAG 452
QY 446 CGACAGACCCACTCCAGTGTGGTGTCCACGTTGTGAGAGTGACCAAGAGAAATGAGG 505
Db 453 CGACTGACACCACTCCAGCGTGGTGTCCATGTATGTGAGAGTGATCAAGAGAGGAGG 512
QY 506 AGNAGAGGAGATGGAGAGATGAAGAGACCAAGCCAAATCATCCAGACACGGA 565
Db 513 AGAAGAGGAA--ATGGAAGAATGAGGAGACCTAAGCCAAATTTATCCAGACCAGGA 569
QY 566 GACCGAGTACACACCGATCCACTTACCTGA 597
Db 570 GCGCGAGTACCGCCGATCCACTCAGCTGA 601

RESULT 13

US-09-724-676A-12745

; Sequence 12745, Application US/09724676A

; GENERAL INFORMATION:

; APPLICANT: Compugen LTD

; TITLE OF INVENTION: Variants of alternative splicing

; FILE REFERENCE: 129181.4 Compugen

; CURRENT APPLICATION NUMBER: US/09/724,676A

; CURRENT FILING DATE: 2000-11-28

; NUMBER OF SEQ ID NOS: 97222

; SOFTWARE: Patent in version 3.2

; SEQ ID NO 12745

; LENGTH: 1876

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-724-676A-12745

Query Match 67.3%; Score 401.6; DB 5; Length 1876;

Best Local Similarity 87.9%; Pred. No. 5.3e-108;

Matches 450; Conservative 0; Mismatches 59; Indels 3; Gaps 1;

QY 86 GGGCCAAATTTGAATCCCTTCAGAACATATGACAAAGACACCACTTCCAGTATTTTA 145
Db 93 GGGCCAAATTTGAATCCCTTCAGAACATATGACAAAGACATCACTTTCAGTATTTTA 152
QY 146 AGAGCTTCAAACTGTCCGATAACTTCAGCAACCCCTTATCTGCGCGGATGCCAGC 205
Db 153 AGAGCTTCAAACTGTCCGATAACTTCAGCAACCCCTTCTCCGAGCAGATGCCAGC 212
QY 206 TGGCGTGCACAAAGACCGAGTTCTCTGGGAAGGAAATGAAGTTGATTTTGTCTCAGACTT 265
Db 213 TCAGCTGCATAGACTGAGTTCTCTGGGAAGGAAATGAAGTTATATTTTGTCTCAGACTT 272
QY 266 TACACATAGGAAGTTACACCTGGCTCGGCCAATCCGCAACAGTTCTCTCATCTCCC 325
Db 273 TACACATAGGAAGTTACACCTGGCTCGGCCAATCCGCAACAGTTCTCTCATCTCCC 332
QY 326 CTCCGGCTCTCTCCGTTGGCTGGAAACAAAGTAGAGATGCCACCCCGTGCATAAAT 385
Db 333 CTCCGGCTCTCTCCGTTGGCTGGAAACAAAGTAGAGATGCCACCCCGTGCATAAAT 392
QY 386 ACGATCTTTATATGTCATCTCCAGCTGGGGCCAGAGAGATGAATGCTGCAG 445
Db 393 ATGATCTCTTATATGTCATCTCCAGCTGGGGCCAGAGAGATGAATGCTGCAG 452

QY 446 CGACAGACCCACTCCAGTGTGGTGTCCACGTTGTGAGAGTGACCAAGAGAAATGAGG 505
Db 453 CGACTGACACCACTCCAGCGTGGTGTCCATGTATGTGAGAGTGATCAAGAGAGGAGG 512
QY 506 AGAAGAGGAAAGATGGAGAGATGAAGAGACCCAAAGCCAAATCATCCAGACACGGA 565
Db 513 AGAAGAGGAA--ATGGAAGAATGAGGAGACCTAAGCCAAATTTATCCAGACCAGGA 569
QY 566 GACCGAGTACACACCGATCCACTTACCTGA 597
Db 570 GCGCGAGTACCGCGATCCACTCAGCTGA 601

RESULT 14

US-09-724-676-12749

; Sequence 12749, Application US/09724676

; GENERAL INFORMATION:

; APPLICANT: Compugen LTD

; TITLE OF INVENTION: Variants of alternative splicing

; FILE REFERENCE: 129181.4 Compugen

; CURRENT APPLICATION NUMBER: US/09/724,676

; CURRENT FILING DATE: 2000-11-28

; NUMBER OF SEQ ID NOS: 97222

; SOFTWARE: Patent in version 3.2

; SEQ ID NO 12749

; LENGTH: 911

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (11)..(11)

; OTHER INFORMATION: n is a,c,g, or t

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (18)..(18)

; OTHER INFORMATION: n is a,c,g, or t

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (23)..(23)

; OTHER INFORMATION: n is a,c,g, or t

US-09-724-676-12749

Query Match 58.7%; Score 350.6; DB 5; Length 911;

Best Local Similarity 88.6%; Pred. No. 3.8e-93;

Matches 380; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1 ATGGATTTTAGGAGCTTTAGCTTACAAATTTTAGCTCCCTGATTGCTTGTGTCGAACGAT 60
Db 63 ATGCATTTTAGAAACTTTTAACTTACAGTTTGTAGCTCCCTGATTGCTGTCGCAACAGT 122
QY 61 GATGCTTTCAGCGAAGTGAGACCAAGGCCAAATTTGAATCCCTCTTCAGAACATATGAC 120
Db 123 GATATCTTCAGCGAAGTGAAACCAAGGCCAAATTTGAGTCCCTCTTTAGGACGTTATGAC 182
QY 121 AAGGACACACCTTCCAGTATTTTAAAGAGCTTCAAACTGTCGCGATAAACTTCAGCAAC 180
Db 183 AAGGACATCACTTTCAGTATTTTAAAGAGCTTCAAACTGAGTCAAGTAACTTCAGCAAC 242
QY 181 CCCTTATCTGACGCGATGTCAGGCTGCGGTGCAACAAGCCGAGTTCTCTGGGGAAGGAA 240
Db 243. CCCTTCTCCGAGCAGATGCCAGGCTCCAGCTGCATAGACTGAGTTTCTGGGAAGGAA 302
QY 241 ATGAAGTTGATTTTGTCTCAGACTTTTACACATAGGAAGTTTACACCTGGCTCCGCCCAAT 300
Db 303 ATGAAGTTTATTTTGTCTCAGACTTTTACACATAGGAAGTTTACACCTGGCTCCGCCCAAT 362
QY 301 CCCGACAAACAGTTTCTCTCATCTCCCTCCGCGCTCTCTCTCCGCTTGGCTGGAACAAGTA 360
Db 363 CCAGACAGCAGTTTCTGATCTCCCTCCGCGCTCTCTCCGCGCTTGGGATGGAACAAGTG 422
QY 361 GAAAGTGCACCCCGCTCATAAATTCGATCTTTTATATGCTATCTCCAGCTCGGGCCA 420

Db 423 GAAGATGGACCCCACTGTAATACTATGATCTCTATATGCACTCCAAAGCTGGGGCCA 482

QY 421 GGAGGAAG 429

Db 483 GGTGAGCAG 491

Search completed: December 14, 2002, 19:59:19
Job time : 36.1761 secs

RESULT 15
US-09-724-676A-12749

/ Sequence 12749, Application US/09724676A

/ GENERAL INFORMATION:

/ APPLICANT: Compugen LTD

/ TITLE OF INVENTION: Variants of alternative splicing

/ FILE REFERENCE: 129181.4 Compugen

/ CURRENT APPLICATION NUMBER: US/09/724,676A

/ NUMBER OF SEQ ID NOS: 97222

/ SOFTWARE: PatentIn version 3.2

/ SEQ ID NO 12749

/ LENGTH: 911

/ TYPE: DNA

/ ORGANISM: Homo sapiens

/ FEATURE:

/ NAME/KEY: misc_feature

/ LOCATION: (11)..(11)

/ OTHER INFORMATION: n is a,c,g, or t

/ FEATURE:

/ NAME/KEY: misc_feature

/ LOCATION: (18)..(18)

/ OTHER INFORMATION: n is a,c,g, or t

/ FEATURE:

/ NAME/KEY: misc_feature

/ LOCATION: (23)..(23)

/ OTHER INFORMATION: n is a,c,g, or t

/ US-09-724-676A-12749

Query Match

Best Local Similarity 58.7%; Score 350.6; DB 5; Length 911;

Matches 380; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1 ATGGATTTTAAAGGACTTTAGCTACATTTTAACTTCCCTGATTCCTGTGTGGCAACGAT 60

Db 63 ATGCATTTTAAAGAACTTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTT 122

QY 61 GATGCTTCAAGGAAAGTGAGCCAGGCGCAATTGAATCCCTTTCAGAAACATATGAC 120

Db 123 GATATCTTCAGCAAGTGAAGTGAAGCCAGGCGCAATTGAATCCCTTTCAGAAACATATGAC 182

QY 121 AAGGACACCACTTCCAGTATTTTAAAGCTTCAAGCTTCCGATTAACCTTCAGCAAC 180

Db 183 AAGGACATCACTTTCAGTATTTTAAAGCTTCAAGCTTCCGATTAACCTTCAGCAAC 242

QY 181 CCTTATCTGCAAGCCGATGCGAGCTGCGCTCCACAAAGCCGATTAACCTTCAGCAAC 240

Db 243 CCTTATCTGCAAGCCGATGCGAGCTGCGCTCCACAAAGCCGATTAACCTTCAGCAAC 302

QY 241 ATGAAGTGTATTTTGTCTGAGCTTACATAGAAATTCACACCTGCTCCGCCCAAT 300

Db 303 ATGAAGTGTATTTTGTCTGAGCTTACATAGAAATTCACACCTGCTCCGCCCAAT 362

QY 301 CCCGACAAACAGTTCTCTCATCTCCCTCCGCTTCTCCCTGCTGGGAAACAAAGTA 360

Db 363 CCAGACAAAGAGTTTGTGATCTCCCTCCGCTTCTCCCTGCTGGGAAACAAAGTA 422

QY 361 GAAGATGCCACCCCGTCATTAATTAAGATCTTTTAAATGCAATTCCTCAAGCTGGGGCCA 420

Db 423 GAAGATGCCACCCCGTCATTAATTAAGATCTTTTAAATGCAATTCCTCAAGCTGGGGCCA 482

QY 421 GGAGAGAAG 429

Db 483 GGTGAGCAG 491

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	2331	100.0		2331	10	US-09-782-953-11	Sequence 11, Appl
2	2329.4	99.9		2355	12	US-10-044-090-255	Sequence 255, Appl
3	2109	90.5		2212	10	US-09-782-953-14	Sequence 14, Appl
4	2101	90.1		2173	10	US-09-880-107-3340	Sequence 3340, Appl
5	2094.2	89.8		2358	10	US-09-925-302-347	Sequence 347, Appl
6	486	20.8		486	10	US-09-864-761-809	Sequence 809, Appl
7	470.6	20.2		597	10	US-09-782-953-5	Sequence 5, Appl
8	413.8	17.8		599	10	US-09-782-953-1	Sequence 1, Appl
9	403.8	17.3		597	10	US-09-782-953-2	Sequence 2, Appl
10	365.2	15.7		385	10	US-09-778-320-71	Sequence 71, Appl
11	365.2	15.7		385	10	US-09-910-689-71	Sequence 71, Appl
12	365.2	15.7		385	12	US-10-010-742-71	Sequence 71, Appl
13	246.8	10.6		251	10	US-09-745-288-83	Sequence 83, Appl
14	230.2	9.9		473	10	US-09-960-352-6648	Sequence 6648, Appl
15	194.4	8.3		3184	10	US-09-854-456-497	Sequence 497, Appl
16	194.4	8.3		3184	10	US-09-782-953-17	Sequence 17, Appl
17	193.2	8.1		594	10	US-09-782-953-8	Sequence 8, Appl
18	189	8.1		720	10	US-09-782-953-23	Sequence 23, Appl
19	189	8.1		828	10	US-09-782-953-20	Sequence 20, Appl

QY 181 TGATTGCCCTGTGTGGCA

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

QY 181 TGATTGCCCTGTGTGGCAACAGTGATATCTTCAGCGAAAGTGAAACCAGGGCCAAATTTTG 240

Db 181 TGATGGCCGTGTGGCAACAGTGAATCTTCAGCGAAAGTGAACCGAGGCCAAATTTG 240
Qy 241 AGTCCCTCTTTAGACGTATGACAGACATCACTTTCACTATTTTAAGAGCTTCAAC 300
Db 241 AGTCCCTCTTTAGACGTATGACAGACATCACTTTCACTATTTTAAGAGCTTCAAC 300
Qy 301 GAGTCAGAAATTAATCTTCAAGCAACCCCTTCTCCGACGAGATGCCAGCTCCAGCTGCA 360
Db 301 GAGTCAGAAATTAATCTTCAAGCAACCCCTTCTCCGACGAGATGCCAGCTCCAGCTGCA 360
Qy 361 AGACTGAGTTTCTGGGAAAGAAATGAAGTTATTTTGTCTGACAGCTTACACATAGGAA 420
Db 361 AGACTGAGTTTCTGGGAAAGAAATGAAGTTATTTTGTCTGACAGCTTACACATAGGAA 420
Qy 421 GCTCAACCTGGCTCCGCCAAATCCAGACAGCAAGTTTGTATCTTCCCTCCGCTCTC 480
Db 421 GCTCAACCTGGCTCCGCCAAATCCAGACAGCAAGTTTGTATCTTCCCTCCGCTCTC 480
Qy 481 CCCCAGTGGGATGGAACAAAGTGAAGATGCAAGCCAGTCAATAATGATCTTAT 540
Db 481 CCCCAGTGGGATGGAACAAAGTGAAGATGCAAGCCAGTCAATAATGATCTTAT 540
Qy 541 ATGCCATCTCCAAAGCTGGGGCCAGGGGAAAGTATGAATTGACAGCGACGACCA 600
Db 541 ATGCCATCTCCAAAGCTGGGGCCAGGGGAAAGTATGAATTGACAGCGACGACCA 600
Qy 601 CTCCAGCGTGGTGTGCTCATGTATGTGAGAGTATCAAGAAAGGAAAGGAA 660
Db 601 CTCCAGCGTGGTGTGCTCATGTATGTGAGAGTATCAAGAAAGGAAAGGAA 660
Qy 661 TGGAAAGAAATGAGGAAACCTTAAGCCAAATTAATCCAGACGAGGCGGAGTACGC 720
Db 661 TGGAAAGAAATGAGGAAACCTTAAGCCAAATTAATCCAGACGAGGCGGAGTACGC 720
Qy 721 CGATCCAGCTCAAGCTGACAGTGGACGAGGAGGAGCATTCCTCAATCACTACGGA 780
Db 721 CGATCCAGCTCAAGCTGACAGTGGACGAGGAGGAGCATTCCTCAATCACTACGGA 780
Qy 781 GGAATCTTTTACTGTGAGAGTGGCTGTCAGCACTTCTTGGAGGTGGACCGCAATGC 840
Db 781 GGAATCTTTTACTGTGAGAGTGGCTGTCAGCACTTCTTGGAGGTGGACCGCAATGC 840
Qy 841 GGGTGGCAGAAATCCAGATCATGTGCTCAGAAAGAAATCAAGAGCGCTGCCCTGTT 900
Db 841 GGGTGGCAGAAATCCAGATCATGTGCTCAGAAAGAAATCAAGAGCGCTGCCCTGTT 900
Qy 901 CTAAATCTGACACCAAGTTACTGTTCATGACACCGGAAATGAATTGAGCACTGA 960
Db 901 CTAAATCTGACACCAAGTTACTGTTCATGACACCGGAAATGAATTGAGCACTGA 960
Qy 961 GTTTGTGTGATGTGACACAGGACATTTGGGACGTCTTGAAGAAACAGATTAAGATGT 1020
Db 961 GTTTGTGTGATGTGACACAGGACATTTGGGACGTCTTGAAGAAACAGATTAAGATGT 1020
Qy 1021 TTTTGTACTGTCTTCTGTAGAGTCTGTCTGTGCAAGGACAGCTTGAATCACTGAG 1080
Db 1021 TTTTGTACTGTCTTCTGTAGAGTCTGTCTGTGCAAGGACAGCTTGAATCACTGAG 1080
Qy 1081 CTGACGAGAGAGCTTCTGTTTCTAAGTGGCTGACAGGGCCACTCTCTACTGATGAA 1140
Db 1081 CTGACGAGAGAGCTTCTGTTTCTAAGTGGCTGACAGGGCCACTCTCTACTGATGAA 1140
Qy 1141 GAGGTACCAAGAAAGCGCTAGTGACAGAGGTTGTGAAGACAGCAGCAATGATGT 1200
Db 1141 GAGGTACCAAGAAAGCGCTAGTGACAGAGGTTGTGAAGACAGCAGCAATGATGT 1200
Qy 1201 GGAATTTGATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1260
Db 1201 GGAATTTGATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1260
Qy 1261 ATTTACAGACTAACTTTGTTGTTATTAAGTTACCGCTTGTCTTTTACATCTTT 1320

Db 1261 ATTTACAGACTAACTTTGTTGTTATTAAGTTACCGCTTGTCTTTTACATCTTT 1320
Qy 1321 GGGAGGCGAGAAAGCGTTTGGAAACGTAATCACTTCCAGATCTGGATTTCTGAC 1380
Db 1321 GGGAGGCGAGAAAGCGTTTGGAAACGTAATCACTTCCAGATCTGGATTTCTGAC 1380
Qy 1381 TCTTTGCAAGCACTTGTCTTGGGAACTCTTCTGGAATGATCACTCAGATCCCA 1440
Db 1381 TCTTTGCAAGCACTTGTCTTGGGAACTCTTCTGGAATGATCACTCAGATCCCA 1440
Qy 1441 ACCGTGCAAGTGTACTTGTGTCTTTGCAAGAAAGTTGATCTGAATTCCTGTAGA 1500
Db 1441 ACCGTGCAAGTGTACTTGTGTCTTTGCAAGAAAGTTGATCTGAATTCCTGTAGA 1500
Qy 1501 ATTTAGCTTATACAAATTCAGAGAAATAGCACTTCACTGCCAATTTTATGAGGAGAAA 1560
Db 1501 ATTTAGCTTATACAAATTCAGAGAAATAGCACTTCACTGCCAATTTTATGAGGAGAAA 1560
Qy 1561 TTTTATGTTAGTGTGGATGCGACCTGATCTGTTGTTCTTTTATGTGTGTT 1620
Db 1561 TTTTATGTTAGTGTGGATGCGACCTGATCTGTTGTTCTTTTATGTGTGTT 1620
Qy 1621 TCTATACATGAATCATGCCAAATCTTTTGGAACTGTGTGTTGATGATGATGATG 1680
Db 1621 TCTATACATGAATCATGCCAAATCTTTTGGAACTGTGTGTTGATGATGATGATG 1680
Qy 1681 TTTTACCCCAAGAGACATCAAGATACATCTTGAATTAAGCTGATATATATAT 1740
Db 1681 TTTTACCCCAAGAGACATCAAGATACATCTTGAATTAAGCTGATATATATATAT 1740
Qy 1741 CCTGTGTACCTTGGTGAAGAAAGTATGAGGAGGAGCTTAAGATATTAATCTTACT 1800
Db 1741 CCTGTGTACCTTGGTGAAGAAAGTATGAGGAGGAGCTTAAGATATTAATCTTACT 1800
Qy 1801 GTGAATCATATGTTGTGAGAAAGCTGTTCCCATGTCTTAACAGGACTTGAATCAAGCA 1860
Db 1801 GTGAATCATATGTTGTGAGAAAGCTGTTCCCATGTCTTAACAGGACTTGAATCAAGCA 1860
Qy 1861 TGTCAAGTGAATAGTATCTGTGGCATATGAGAGGAGTCAAGTGCCTTCCCATTTCA 1920
Db 1861 TGTCAAGTGAATAGTATCTGTGGCATATGAGAGGAGTCAAGTGCCTTCCCATTTCA 1920
Qy 1921 TTCTGATGGAATTTTATCTAGTTAACTTTGTAATTTTTTTCTAAGTGTAAATGCT 1980
Db 1921 TTCTGATGGAATTTTATCTAGTTAACTTTGTAATTTTTTTCTAAGTGTAAATGCT 1980
Qy 1981 ATGTCTGTAATATAGTATTAATTTTGGCTTACATACCGTAACATGTTGTCAATTT 2040
Db 1981 ATGTCTGTAATATAGTATTAATTTTGGCTTACATACCGTAACATGTTGTCAATTT 2040
Qy 2041 TGAATTAATTAATGCAAGTAAATGATGATCTTTGAAATTTGGAAATGATGTTTATTC 2100
Db 2041 TGAATTAATTAATGCAAGTAAATGATGATCTTTGAAATTTGGAAATGATGTTTATTC 2100
Qy 2101 TTTTGAAGAGCAATATGTTTGAATTAATGCTTGAATGTTCAATCAAGAAATGATTTG 2160
Db 2101 TTTTGAAGAGCAATATGTTTGAATTAATGCTTGAATGTTCAATCAAGAAATGATTTG 2160
Qy 2161 AACGTTCTCAAACTGTTTACGTAATGTTGTAAGAGGAGCGGTTTGGAGAGACCAT 2220
Db 2161 AACGTTCTCAAACTGTTTACGTAATGTTGTAAGAGGAGCGGTTTGGAGAGACCAT 2220
Qy 2221 TGCATGCTGTCCAAAGTGTCTTGTAAAGTCTTTTAACTGAGAGGCTTAACCTCAAA 2280
Db 2221 TGCATGCTGTCCAAAGTGTCTTGTAAAGTCTTTTAACTGAGAGGCTTAACCTCAAA 2280
Qy 2281 ATACTTTTAACTGATCTTAATTAATTAAGGACAGATAGTCTCTTAC 2331
Db 2281 ATACTTTTAACTGATCTTAATTAATTAAGGACAGATAGTCTCTTAC 2331

; Sequence 255, Application US/10044090
; Patent No. US20020137081A1
; GENERAL INFORMATION:
; APPLICANT: Olga Bandman
; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
; FILE REFERENCE: PA-0028 US
; CURRENT APPLICATION NUMBER: US/10/044,090
; CURRENT FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: PERL Program
; SEQ ID NO 255
; LENGTH: 2355
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020137081A1 5511889CB1
US-10-044-090-255

Query Match 99.9%; Score 2329.4; DB 12; Length 2355;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2330; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TTTTTCCTCCAGGAGTGGGCTGGCCCTTACTGCTTTATAGCACCAGCTCAAGA 60
Db TTTTTCCTCCAGGAGTGGGCTGGCCCTTACTGCTTTATAGCACCAGCTCAAGA 67

Qy 61 AGGAACCTACAGCTCTTGGAAAGAAATCTCACTAGGGGCTTGAAGCTGGGTCTGTAG 120
Db AGGAACCTACAGCTCTTGGAAAGAAATCTCACTAGGGGCTTGAAGCTGGGTCTGTAG 127

Qy 121 CGCTTTCACTCTAGAAAGCAAGTGCATTTTAAAGAACTTTAACTACAGTTTGTAGTCCC 180
Db CGCTTTCACTCTAGAAAGCAAGTGCATTTTAAAGAACTTTAACTACAGTTTGTAGTCCC 187

Qy 181 TGAATGCTGTGTGGCAACAGTGCATATCTTCAGCGAAAGTGAACCCAGGCGCAAAATTGG 240
Db TGAATGCTGTGTGGCAACAGTGCATATCTTCAGCGAAAGTGAACCCAGGCGCAAAATTGG 247

Qy 241 AGTCCCTCTTTAGACGTATGACAAGGACATCACTTTTCACTTCACTTCACTTCACTTCACT 300
Db AGTCCCTCTTTAGACGTATGACAAGGACATCACTTTTCACTTCACTTCACTTCACTTCACT 307

Qy 301 GAGTCAGAAATAAATCTTCAGCAACCCCTTCTCCGACGAGATGCCAGGCTCCAGCTGCATA 360
Db GAGTCAGAAATAAATCTTCAGCAACCCCTTCTCCGACGAGATGCCAGGCTCCAGCTGCATA 367

Qy 361 AGACTGAGTTTCTGGGAAAGAAATGAAGTTATATTTTGTCTCAGACCTTACACATAGGAA 420
Db AGACTGAGTTTCTGGGAAAGAAATGAAGTTATATTTTGTCTCAGACCTTACACATAGGAA 427

Qy 421 GCTCACACCTGGCTCCGCCAAATCCAGCAAGCAGTTTCTGATCTCCCTCCGCGCTCTC 480
Db GCTCACACCTGGCTCCGCCAAATCCAGCAAGCAGTTTCTGATCTCCCTCCGCGCTCTC 487

Qy 481 CGCCAGTGGGATGGAACAAAGTGGAAAGATCGGACCCAGCTCAAACTATGATCTCTTAT 540
Db CGCCAGTGGGATGGAACAAAGTGGAAAGATCGGACCCAGCTCAAACTATGATCTCTTAT 547

Qy 541 ATGCCATCTCAAGCTGGGCGCAGGGGAAAGATGAAATGCAAGCAGGCTGACACCA 600
Db ATGCCATCTCAAGCTGGGCGCAGGGGAAAGATGAAATGCAAGCAGGCTGACACCA 607

Qy 601 CTCCAGCGTGGTGGTCCATGTATGTAGAGTGATCAAGAGAGGAGGAGGAGGAGAA 660
Db CTCCAGCGTGGTGGTCCATGTATGTAGAGTGATCAAGAGAGGAGGAGGAGGAGGAGAA 667

Qy 661 TGAAGAAATAGGAGACCTTAAGCAAAATATTCAGACACGAGGCGGAGGTACAGC 720
Db TGAAGAAATAGGAGACCTTAAGCAAAATATTCAGACACGAGGCGGAGGTACAGC 727

Qy 721 CGATCCACCTCAGCTGAATCGGCAAGGAGCGCATTCCTCAATCATATCTACGGGA 780
Db CGATCCACCTCAGCTGAATCGGCAAGGAGCGCATTCCTCAATCATATCTACGGGA 787

Db 728 CGATCCACCTCAGCTGAATCGGCAAGGAGCGCATTCCTCAATCATATCTACGGGA 787
Qy 781 GGAATCTTTTACTGTGGAGTGCTGGTCAACGACTTCTTCGGAGGTGGCAGCGGATCG 840
Db 788 GGAATCTTTTACTGTGGAGTGCTGGTCAACGACTTCTTCGGAGGTGGCAGCGGATCG 847
Qy 841 GGGTGGCAGAAATCCAGTTTCATGTTGCTCAGAAAGAAATCAAGGCGGTGCCCTTGT 900
Db 848 GGGTGGCAGAAATCCAGTTTCATGTTGCTCAGAAAGAAATCAAGGCGGTGCCCTTGT 907
Qy 901 CTAATGCTGCACACCACTTACTGTTTCATGCGCACCCGGGAATGACTTGGGCCCACTCA 960
Db 908 CTAATGCTGCACACCACTTACTGTTTCATGCGCACCCGGGAATGACTTGGGCCCACTCA 967
Qy 961 GTTTGTGGTATCGCAACAGGACATTTGGGACTGTCTTGAGAAACAGATAATGATAGTG 1020
Db 968 GTTTGTGGTATCGCAACAGGACATTTGGGACTGTCTTGAGAAACAGATAATGATAGTG 1027
Qy 1021 TTTTGTACTTGTCTTCTGTTAGTGTCTGCTGCTGCGCAAGGCGGAGTTGATCAGTGAG 1080
Db 1028 TTTTGTACTTGTCTTCTGTTAGTGTCTGCTGCTGCGCAAGGCGGAGTTGATCAGTGAG 1087
Qy 1081 CTCAGGAGAGCTTCTGTTTCTAAGTGGCTGCGAGGGGCGCACTCTCTACTGGTAGGAA 1140
Db 1088 CTCAGGAGAGCTTCTGTTTCTAAGTGGCTGCGAGGGGCGCACTCTCTACTGGTAGGAA 1147
Qy 1141 GAGGTACACACAGGAGCGCTTGTGCGAGAGGTTGTGAAACAGCAATGCAATGT 1200
Db 1148 GAGGTACACACAGGAGCGCTTGTGCGAGAGGTTGTGAAACAGCAATGCAATGT 1207
Qy 1201 GGAATTTGTAGCGTTTCTTCTCCCTCATGTTCTCATGTTTGTGCGATGATATACTG 1260
Db 1208 GGAATTTGTAGCGTTTCTTCTCCCTCATGTTCTCATGTTTGTGCGATGATATACTG 1267
Qy 1261 ATTTACAGACTTAACCTTTGTTTGTATATAAGTTTACACCGTTTGTGTTTACATCTTT 1320
Db 1268 ATTTACAGACTTAACCTTTGTTTGTATATAAGTTTACACCGTTTGTGTTTACATCTTT 1327
Qy 1321 GGAAGCCAGGAGGAGGTTTGGAAACGATATCACCTTCCAGATCTTCGGATCTCGAC 1380
Db 1328 GGAAGCCAGGAGGAGGTTTGGAAACGATATCACCTTCCAGATCTTCGGATCTCGAC 1387
Qy 1381 TCTTTGCAACAGCACTTCTGTTCCGGAACCTCTTCTGGAATGCACTTCTCAGATCTCCAC 1440
Db 1388 TCTTTGCAACAGCACTTCTGTTCCGGAACCTCTTCTGGAATGCACTTCTCAGATCTCCAC 1447
Qy 1441 ACCGTGCAAGCTGTAACTTGTGCTTTTGCAGAAAGTGTGATCTGAAATTCCTCTGTAGA 1500
Db 1448 ACCGTGCAAGCTGTAACTTGTGCTTTTGCAGAAAGTGTGATCTGAAATTCCTCTGTAGA 1507
Qy 1501 ATTTAGCTTATACAAATTCAGAGAAATAGCAGTTTCCAGTCCCACTTTTAGTGGGTGAGAA 1560
Db 1508 ATTTAGCTTATACAAATTCAGAGAAATAGCAGTTTCCAGTCCCACTTTTAGTGGGTGAGAA 1567
Qy 1561 TTTTAGTTTAGGTGTTTGGGATCGGACCTCAGTTTCTGTTGTTTCTTTTATGCTGGT 1620
Db 1568 TTTTAGTTTAGGTGTTTGGGATCGGACCTCAGTTTCTGTTGTTTCTTTTATGCTGGT 1627
Qy 1621 TCTATACATGAATCATAGCCCAAACTTTTGGAAACTGTTTGGAACTGTTGTTGAGTGTGTT 1680
Db 1628 TCTATACATGAATCATAGCCCAAACTTTTGGAAACTGTTTGGAACTGTTGTTGAGTGTGTT 1687
Qy 1681 TTTTACCCCGAGAGACATCAAGATACACTTGTAAATTAAGCTGTATAGCATATATTCATA 1740
Db 1688 TTTTACCCCGAGAGACATCAAGATACACTTGTAAATTAAGCTGTATAGCATATATTCATA 1747
Qy 1741 CTTGTTGTACATTTGGGTGAAAAGTATGGCAGTGGGAGACTAAGATGTAATTAACCTACCT 1800
Db 1748 CTTGTTGTACATTTGGGTGAAAAGTATGGCAGTGGGAGACTAAGATGTAATTAACCTACCT 1807
Qy 1801 GTGAATCATATGTTGTAAGGAAAGCTGTTCCCATGCTTAACAGGACTTGAATTCAGAGCA 1860
Db 1808 GTGAATCATATGTTGTAAGGAAAGCTGTTCCCATGCTTAACAGGACTTGAATTCAGAGCA 1867


```

Db 778 CAATCAGAGTTTGTGATGCAAGACATTGGAGTGTCTGAGAAAACAGAT 837
Qy 1011 AATGATAGTGTGTACTGTTCTTTCTGTGTAGGTTCTGTGTGCAAGGAGGTT 1070
Db 838 AATGATAGTGTGTACTGTTCTTTCTGTGTAGGTTCTGTGTGCAAGGAGGTT 897
Qy 1071 GATCAGTAGCTCAGAGAGAGCTTCCGTGTTCTTAAGGGCCCTGCAAGGAGGTTCT 1130
Db 898 GATCAGTAGCTCAGAGAGAGCTTCCGTGTTCTTAAGGGCCCTGCAAGGAGGTTCT 957
Qy 1131 CTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1190
Db 958 CTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1017
Qy 1191 AATGCAATGTGAAATTTAGCGTTTCTTTCTTCCCTCATGTCTCATGTTTGTGAG 1250
Db 1018 AATGCAATGTGAAATTTAGCGTTTCTTTCTTCCCTCATGTCTCATGTTTGTGAG 1077
Qy 1251 TATATTACTGATTTCAAGACTAAGCTTGTGATATTAAGTTACACCGTTGTGTTT 1310
Db 1078 TATATTACTGATTTCAAGACTAAGCTTGTGATATTAAGTTACACCGTTGTGTTT 1137
Qy 1311 TACATCTTTTGGAGAGCCAGAGAAAGCGTTTGGAGAAAGATATCACTTTCCAGATTCTG 1370
Db 1138 TACATCTTTTGGAGAGCCAGAGAAAGCGTTTGGAGAAAGATATCACTTTCCAGATTCTG 1197
Qy 1371 GATTCTGACTCTTTTGCACAGACACTTGTGTGGAGAGCTTCCCTGGAGATGATTCAGTC 1430
Db 1198 GATTCTGACTCTTTTGCACAGACACTTGTGTGGAGAGCTTCCCTGGAGATGATTCAGTC 1257
Qy 1431 AGCATCCCAACCGTGCACAGCTGTAACTGTGCTTTTGGAGAAAGATGATTCAGTAAT 1490
Db 1258 AGCATCCCAACCGTGCACAGCTGTAACTGTGCTTTTGGAGAAAGATGATTCAGTAAT 1317
Qy 1491 CCTCTGTAGAAATTTAGCTTATACAAATTCAGAGAAATGACGTTTCACTGCCAATTACT 1550
Db 1318 CCTCTGTAGAAATTTAGCTTATACAAATTCAGAGAAATGACGTTTCACTGCCAATTACT 1377
Qy 1551 GGGTGAGAAATTTTATGTTAGTGTGTGGAGTGGAGCTGAGTTGCTGTTCTTTTA 1610
Db 1378 GGGTGAGAAATTTTATGTTAGTGTGTGGAGTGGAGCTGAGTTGCTGTTCTTTTA 1437
Qy 1611 TGTGTGCTTCTATACATGATCATGACCAAAACTTTTGGAGAACTGTGTTGAGA 1670
Db 1438 TGTGTGCTTCTATACATGATCATGACCAAAACTTTTGGAGAACTGTGTTGAGA 1497
Qy 1671 TAGTTGCTTCTTACCCCAAGAGACATCAAGATACACTTGTAAATTAAGCTGATAGCA 1730
Db 1498 TAGTTGCTTCTTACCCCAAGAGACATCAAGATACACTTGTAAATTAAGCTGATAGCA 1557
Qy 1731 TATATTACTGCTTGTACACTTGGTGAAGAGATGAGAGTGGAGACTAGATATAT 1790
Db 1558 TATATTACTGCTTGTACACTTGGTGAAGAGATGAGAGTGGAGACTAGATATAT 1617
Qy 1791 TAACTTACTGTGATCATATGTTGTAGAGAAAGCTTGTCCATGTCTAAACAGACTTGA 1850
Db 1618 TAACTTACTGTGATCATATGTTGTAGAGAAAGCTTGTCCATGTCTAAACAGACTTGA 1677
Qy 1851 ATTCAAGCATGTCAAGTGTAGTATATCTGTGCGATATGAGAGGATGAGAGTGTCTT 1910
Db 1678 ATTCAAGCATGTCAAGTGTAGTATATCTGTGCGATATGAGAGGATGAGAGTGTCTT 1737
Qy 1911 TCCCATTCATTCCTGATGAGATGTTATACTAGTTAACTTGTATTTTCTTGT 1970
Db 1738 TCCCATTCATTCCTGATGAGATGTTATACTAGTTAACTTGTATTTTCTTGT 1797
Qy 1971 TGTATGTATGTCTGTGTAATAGTATTAATTTTGGCTTAAATACCTTAACATG 2030
Db 1798 TGTATGTATGTCTGTGTAATAGTATTAATTTTGGCTTAAATACCTTAACATG 1857
Qy 2031 TTGTGATTTGAAATCTTAATGCAAGTAAACAATGATCTTTGAGAAATTTGAGAT 2090

```

```

Db 1858 TTGTGATTTTGAATACTTAATGCCAGTAAACATGATGCTTTGAAATTTGAGAT 1917
Qy 2091 GGTATTATCTTTGAGAGCAATATGTTGATTAATATGCTTTGATTTGATATCAAG 2150
Db 1918 GGTATTATCTTTGAGAGCAATATGTTGATTAATATGCTTTGATTTGATATCAAG 1977
Qy 2151 AAATTGATGAGCTTCTCAACACCTGTTTACGTTACTGTTGTAAGGAGAGCGGTTTG 2210
Db 1978 AAATTGATGAGCTTCTCAACACCTGTTTACGTTACTGTTGTAAGGAGAGCGGTTTG 2037
Qy 2211 GAGAGACATTTGATGCTGCTCCAGAGTGTCTTGTAAAGTCTTTTAACTGAGAGGC 2270
Db 2038 GAGAGACATTTGATGCTGCTCCAGAGTGTCTTGTAAAGTCTTTTAACTGAGAGGC 2097
Qy 2271 TAACTTCAAAATATCTTTTAACTGATTTCTATTAATTAATGAGGACAGATATCTCTTA 2330
Db 2098 TAACTTCAAAATATCTTTTAACTGATTTCTATTAATTAATGAGGACAGATATCTCTTA 2157
Qy 2331 C 2331
Db 2158 C 2158

```

```

RESULT 5
US-09-925-302-347
/ Sequence 347, Application US/09925302
/ Patent No. US20020044941A1
/ GENERAL INFORMATION:
/ APPLICANT: Rosen et al.
/ TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
/ FILE REFERENCE: P0104
/ CURRENT APPLICATION NUMBER: US/09/925.302
/ PRIOR FILING DATE: 2001-08-10
/ PRIOR APPLICATION NUMBER: PCT/US00/05918
/ PRIOR FILING DATE: 2000-03-08
/ PRIOR APPLICATION NUMBER: 60/124,270
/ NUMBER OF SEQ ID NOS: 896
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 347
/ LENGTH: 2358
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-302-347

```

```

Query Match 89.8%; Score 2094.2; DB 10; Length 2358;
Best Local Similarity 97.3%; Pred. No. 0;
Matches 2137; Conservative 3; Mismatches 56; Indels 1; Gaps 1;

```

```

Qy 135 GAAAGCAAGATGATTTTGAAGAACTTTAAGTACAGTTTAACTGCTGATTTGGCTGTG 194
Db 105 GACTCGAGATGAGAGAGGTGACCTGACAGAGCTGCCAGGCGACCATGCTGTAC 164
Qy 195 GCAAACAGTATATCTTACGCGAGAGAGTAAACAGAGGCGCAATTTTGAATCCTCTTAGG 254
Db 165 CTGAGACCGCGGTGTGTGTGTGAGAGGCGGTGCGGGCAATTTGAATGCTCTTAGG 224
Qy 255 ACGTATGACAGAGATCACTTTTCAATTTTAAAGCTTAAACGAGTCAAGATTAAC 314
Db 225 ACGTATGACAGAGATCACTTTTCAATTTTAAAGCTTAAACGAGTCAAGATTAAC 284
Qy 315 TTTCAGAACCCCTTCTCCGACAGATGCGAGGCTCCAGCTCATTAAGTATTTCTG 374
Db 285 TTTCAGAACCCCTTCTCCGACAGATGCGAGGCTCCAGCTCATTAAGTATTTCTG 344
Qy 375 GGAAGAGATGAGTATATTTTGTCTAGACCTTACATAGAGAGCTCACACCTGCT 434
Db 345 GGAAGAGATGAGTATATTTTGTCTAGACCTTACATAGAGAGCTCACACCTGCT 404
Qy 435 CGGCCAAATCCAGACAGAGATTTCTGATCTCCCTCCGCTTCCGCACTGGAGTGG 494
Db 405 CCG-CAAAATCCAGACAGAGATTTCTGATCTCCCTCCGCTTCCGCACTGGAGTGG 463

```

QY 495. AAACAAGTGAAGATCGACCCAGTCATATAAATATGATCTCTTATATGCCATCTCCAAG 554
Db 464 AAACAAGTGAAGATCGACCCAGTCATATAAATATGATCTCTTATATGCCATCTCCAAG 523
QY 555 CTGGGCCAGGGGAAAGATGATGAATTCAGCAGCAGGACTGACACCACTCCACAGCGTGGTG 614
Db 524 CTGGGCCAGGGGAAAGATGATGAATTCAGCAGCAGGACTGACACCACTCCACAGCGTGGTG 583
QY 615 GTCCATGTATGTAGAGTGTATCAAGAGAGAGGAGGAAATGGAAGAATGAGG 674
Db 584 GTCCATGTATGTAGAGTGTATCAAGAGAGAGGAGGAAATGGAAGAATGAGG 643
QY 675 AGACCTTAAGCCAAAATTTATCCAGACACGAGGCGGAGTACACGCCGATCCACTCAGC 734
Db 644 AGACCTTAAGCCAAAATTTATCCAGACACGAGGCGGAGTACACGCCGATCCACTCAGC 703
QY 735 TGAACCTGGCAGCGACGAGGACGATTCCAAATCATATCTACCGGAGGAATCTTTTACTG 794
Db 704 TGAACCTGGCAGCGACGAGGACGATTCCAAATCATATCTACCGGAGGAATCTTTTACTG 763
QY 795 TGGAGTGGCTGGTCAACCACTCTTCGGAGTGGGAGCGCGAGATCGGGTGGCAGAAATC 854
Db 764 TGGAGTGGCTGGTCAACCACTCTTCGGAGTGGGAGCGCGAGATCGGGTGGCAGAAATC 823
QY 855 CCAGTTCAATGTTGCTCAGAAGAGAAATCAAGCCGCTGTCCTTCTTAATGCTGCACAC 914
Db 824 CCAGTTCAATGTTGCTCAGAAGAGAAATCAAGCCGCTGTCCTTCTTAATGCTGCACAC 883
QY 915 CAGTTACTGTTTCAATGGCACCAGGAAATGACTTGGCCCAATCACTGAGTTTGTGTATCG 974
Db 884 CAGTTACTGTTTCAATGGCACCAGGAAATGACTTGGCCCAATCACTGAGTTTGTGTATCG 943
QY 975 CACAAGGACATTTGGGACTGCTTTGAGAAAACAGATATGATAGTGTGTTGTAATGTTTC 1034
Db 944 CACAAGGACATTTGGGACTGCTTTGAGAAAACAGATATGATAGTGTGTTGTAATGTTTC 1003
QY 1035 TTTTCTGGTAGTTCTGTCTGTGCAAGGCGCAGTTTCATCAGTGAAGTCTGAGGAGAGCT 1094
Db 1004 TTTTCTGGTAGTTCTGTCTGTGCAAGGCGCAGTTTCATCAGTGAAGTCTGAGGAGAGCT 1063
QY 1095 TCCTGTTTCTTAAGTGGCTCGAGGGCCACTCTCTACTGGTAGAAGAGGTACCAAGCA 1154
Db 1064 TCCTGTTTCTTAAGTGGCTCGAGGGCCACTCTCTACTGGTAGAAGAGGTACCAAGCA 1123
QY 1155 AGCGCCCTAGTGCAGAGAGTGTGAAAACAGCAGCAATGCAATGCGAAATGTAAGCT 1214
Db 1124 AGCGCCCTAGTGCAGAGAGTGTGAAAACAGCAGCAATGCAATGCGAAATGTAAGCT 1183
QY 1215 TTCTTTCTTCCCTCATGTTCTCATGTTTGTGCATGTATATTAATTAATTAATTAATTA 1274
Db 1184 TTCTTTCTTCCCTCATGTTCTCATGTTTGTGCATGTATATTAATTAATTAATTAATTA 1243
QY 1275 CCTTTGTTCTGATATAAAGTTACACCGTTGTTGTTTACATCTTTTGGAAAGCCAGGAA 1334
Db 1244 CCTTTGTTCTGATATAAAGTTACACCGTTGTTGTTTACATCTTTTGGAAAGCCAGGAA 1303
QY 1335 GCGTTTGGAAAACGATATCACTTTCCAGATCTCGGATCTCGACTCTTTGCAACAGCA 1394
Db 1304 GCGTTTGGAAAACGATATCACTTTCCAGATCTCGGATCTCGACTCTTTGCAACAGCA 1363
QY 1395 CTGTGCTGGGAACTCTTCTGGAATGCAATCACTCAGCATCCCAACCGTGCAACGTT 1454
Db 1364 CTGTGCTGGGAACTCTTCTGGAATGCAATCACTCAGCATCCCAACCGTGCAACGTT 1423
QY 1455 AACTGTGCTTTTGGAAAAGAGTTGATCTGAAATCTCTCTGTAGAAATTTAGCTTATACA 1514
Db 1424 AACTGTGCTTTTGGAAAAGAGTTGATCTGAAATCTCTCTGTAGAAATTTAGCTTATACA 1483
QY 1515 ATTACAGAAATAGCAGTTTCACTGCGCAACTTTTAGTGGGTGAGAAATTTAGTTAGTG 1574
Db 1484 ATTACAGAAATAGCAGTTTCACTGCGCAACTTTTAGTGGGTGAGAAATTTAGTTAGTG 1543
QY 1575 TTTGGGATCGGACCTCAGTTTCTGTTGTTTCTTTTATGTTGGTGTCTTCTATACATGAATC 1634

Db 1544 TTTGGGATCGGACCTCAGTTTCTGTTGTTCTTTTATGTTGGTGTCTTATACATGAATC 1603
QY 1635 ATAGCCAAAACCTTTTGGAAACTGTTGGTGTGAGATGTTGTTCTTTTACCCACGAA 1694
Db 1604 ATAGCCAAAACCTTTTGGAAACTGTTGGTGTGAGATGTTGTTCTTTTACCCACGAA 1663
QY 1695 GACATCAAGATACACCTTTGTAATAAGCTGTAGATGATATATTCATACCTGTTGTACACTT 1754
Db 1664 GACATCAAGATACACCTTTGTAATAAGCTGTAGATGATATATTCATACCTGTTGTACACTT 1723
QY 1755 GGGTAAAAAGTATGGCAGTGGGAGACTAAGATGTATTAACTTACCTACCTGTGAATCATATGTT 1814
Db 1724 GGGTAAAAAGTATGGCAGTGGGAGACTAAGATGTATTAACTTACCTACCTGTGAATCATATGTT 1783
QY 1815 GTAGGAAAAGCTGTTCCCATGTCTAACAGGACTTGAATTCAAAGCATGTCAAGTGGATAG 1874
Db 1784 GTAGGAAAAGCTGTTCCCATGTCTAACAGGACTTGAATTCAAAGCATGTCAAGTGGATAG 1843
QY 1875 TAGATCTGGCGATATGAGAGGGATGAGTGCCTTTCCCATTCATTCTCTGATGGAATTT 1934
Db 1844 TAGATCTGGCGATATGAGAGGGATGAGTGCCTTTCCCATTCATTCTCTGATGGAATTT 1903
QY 1935 GTTATCTAGGTAAACATTTGTAATTTTTTCTAGTGTGTAATGTATGTCTGTGTAATA 1994
Db 1904 GTTATCTAGGTAAACATTTGTAATTTTTTCTAGTGTGTAATGTATGTCTGTGTAATA 1963
QY 1995 GGTATTTATTTTGGCTTACAATACCGTAAACCAATGTTTGTCAATTTTGAAGCAAAAT 2054
Db 1964 GGTATTTATTTTGGCTTACAATACCGTAAACCAATGTTTGTCAATTTTGAAGCAAAAT 2023
QY 2055 CCAAGTAAACATGCAATGCTTTGGAAATTTGGAAGATGTTTATCTTTGAGAGCAAAAT 2114
Db 2024 CCAAGTAAACATGCAATGCTTTGGAAATTTGGAAGATGTTTATCTTTGAGAGCAAAAT 2083
QY 2115 ATGTTTGCATTAATGCTTTGATTTGTTTCATATCAAGAAATGATTGAACGTTCTCAAAAC 2174
Db 2084 ATGTTTGCATTAATGCTTTGATTTGTTTCATATCAAGAAATGATTGAACGTTCTCAAAAC 2143
QY 2175 CTGTTTACGGTACTTGTGTAAGAGGAGCGGTTTGGGAGAGACCATTTGCAATGCTGTCCA 2234
Db 2144 CTGTTTACGGTACTTGTGTAAGAGGAGCGGTTTGGGAGAGACCATTTGCAATGCTGTCCA 2203
QY 2235 AGTGTTCCTCTTAAGTGTCTTTAACTGGAGAGGCTAACTCAAAATACCTTTTAACT 2294
Db 2204 AGTGTTCCTCTTAAGTGTCTTTAACTGGAGAGGCTAACTCAAAATATTTTTTTTAACT 2263
QY 2295 TGCATTTCTTAATAAATGAGGACAGTATGCTCTCTTAC 2331
Db 2264 TGCATTTCTTAATAAATGAGGACAGTATGCTCTCTTAC 2300

RESULT 6

US-09-864-761-809/c

; Sequence 809, Application US/09864761

; Patent No. US20020048763A1

; GENERAL INFORMATION:

; APPLICANT: Penn, Sharron G.

; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.

; APPLICANT: Chen, Wensheng

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

; FILE REFERENCE: Aemica-X-1

; CURRENT APPLICATION NUMBER: US/09/864, 761

; PRIOR FILING DATE: 2001-05-23

; PRIOR APPLICATION NUMBER: US 60/180,312

; PRIOR FILING DATE: 2000-02-04

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: US 09/632,366

; PRIOR FILING DATE: 2000-08-03

; PRIOR APPLICATION NUMBER: GB 24263.6


```

/ PRIOR FILING DATE: 2000-10-04
/ PRIOR APPLICATION NUMBER: US 60/236,359
/ PRIOR FILING DATE: 2000-09-27
/ PRIOR APPLICATION NUMBER: PCT/US01/00666
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00667
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00664
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00669
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00665
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00668
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00663
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00662
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00661
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00670
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: US 60/234,687
/ PRIOR FILING DATE: 2000-09-21
/ PRIOR APPLICATION NUMBER: US 09/608,408
/ PRIOR FILING DATE: 2000-06-30
/ PRIOR APPLICATION NUMBER: US 09/774,203
/ PRIOR FILING DATE: 2001-01-29
/ NUMBER OF SEQ ID NOS: 49117
/ SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
/ SEQ ID NO: 809
/ LENGTH: 486
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: MAP TO Ap000169.1
/ OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 4.7
/ OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.6
/ OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.1
/ OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.9
/ OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7
/ OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 6.6
/ OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2
/ OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 8.2
/ OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.3
/ OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.9
/ US-09-864-761-809

Query Match          20.8%; Score 486; DB 10; Length 486;
Best Local Similarity 100.0%; Pred. No. 6,2e-125;
Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

/ PRIOR FILING DATE: 2000-10-04
/ PRIOR APPLICATION NUMBER: US 60/236,359
/ PRIOR FILING DATE: 2000-09-27
/ PRIOR APPLICATION NUMBER: PCT/US01/00666
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00667
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00664
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00669
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00665
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00668
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00663
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00662
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00661
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00670
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: US 60/234,687
/ PRIOR FILING DATE: 2000-09-21
/ PRIOR APPLICATION NUMBER: US 09/608,408
/ PRIOR FILING DATE: 2000-06-30
/ PRIOR APPLICATION NUMBER: US 09/774,203
/ PRIOR FILING DATE: 2001-01-29
/ NUMBER OF SEQ ID NOS: 49117
/ SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
/ SEQ ID NO: 809
/ LENGTH: 486
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: MAP TO Ap000169.1
/ OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 4.7
/ OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.6
/ OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.1
/ OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.9
/ OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7
/ OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 6.6
/ OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2
/ OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 8.2
/ OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.3
/ OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.9
/ US-09-864-761-809

Query Match          20.8%; Score 486; DB 10; Length 486;
Best Local Similarity 100.0%; Pred. No. 6,2e-125;
Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

/ PRIOR FILING DATE: 2000-10-04
/ PRIOR APPLICATION NUMBER: US 60/236,359
/ PRIOR FILING DATE: 2000-09-27
/ PRIOR APPLICATION NUMBER: PCT/US01/00666
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00667
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00664
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00669
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00665
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00668
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00663
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00662
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00661
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00670
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: US 60/234,687
/ PRIOR FILING DATE: 2000-09-21
/ PRIOR APPLICATION NUMBER: US 09/608,408
/ PRIOR FILING DATE: 2000-06-30
/ PRIOR APPLICATION NUMBER: US 09/774,203
/ PRIOR FILING DATE: 2001-01-29
/ NUMBER OF SEQ ID NOS: 49117
/ SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
/ SEQ ID NO: 809
/ LENGTH: 486
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: MAP TO Ap000169.1
/ OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 4.7
/ OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.6
/ OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.1
/ OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.9
/ OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7
/ OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 6.6
/ OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2
/ OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 8.2
/ OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.3
/ OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.9
/ US-09-864-761-809

Query Match          20.8%; Score 486; DB 10; Length 486;
Best Local Similarity 100.0%; Pred. No. 6,2e-125;
Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```


Qy 711 GAGTACAGCCGATCCACTCAGCTGA 737
 Db 571 GAGTACAGCCGATCCACTCAGCTGA 597

RESULT 10

US-09-778-320-71/c
 ; Sequence 71, Application US/09778320
 ; Patent No. US20010034052A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Dillon, Davin C.
 ; APPLICANT: Day, Craig H.
 ; APPLICANT: Jiang, Yugu
 ; APPLICANT: Houghton, Raymond L.
 ; APPLICANT: Mitcham, Jennifer
 ; APPLICANT: Wang, Tongtong
 ; APPLICANT: McNeill, Patricia D.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 ; FILE OF INVENTION: DIAGNOSIS OF BREAST CANCER
 ; FILE REFERENCE: 210121.491C5
 ; CURRENT APPLICATION NUMBER: US/09/778,320
 ; CURRENT FILING DATE: 2001-02-06
 ; NUMBER OF SEQ ID NOS: 301
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 71
 ; LENGTH: 385
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 ; FEATURE:
 ; NAME/KEY: misc.feature
 ; LOCATION: (1)..(385)
 ; OTHER INFORMATION: n = A,T,C or G
 ; US-09-778-320-71

Query Match 15.7%; Score 365.2; DB 10; Length 385;
 Best Local Similarity 98.2%; Pred. No. 1.7e-91;
 Matches 378; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

Qy 752 AGAAGCATTCCTCAATCATCTCAGGAGAGAACTTTTACTGTGAGGTGCTGCTAC 811
 Db 385 AGAGCGATTCCTCAATCATCTCAGGAGAGAACTTTTACTGTGAGGTGCTGCTAC 326
 Qy 812 GACTTCTTCGAGGTGCGAGCGAGAT-CGGGGTGGCAGAAATCCAGTTTATGTTCTC 870
 Db 325 GACTTCTTCGAGGTGCGAGCGAGATCCGGGGTGGCAGAAATCCAGTTTATGTTCTC 266
 Qy 871 AGAAGCATTCCTCAATCATCTCAGGAGAGAACTTTTACTGTGAGGTGCTGCTAC 930
 Db 265 AGAAGCATTCCTCAATCATCTCAGGAGAGAACTTTTACTGTGAGGTGCTGCTAC 206
 Qy 931 CACCGGAAATGACTTGGGCAATCATCTGAGTTTGTGATGCGACAGGATTTGGG 990
 Db 205 CACCGGAAATGACTTGGGCAATCATCTGAGTTTGTGATGCGACAGGATTTGGG 146
 Qy 991 ACTGCTTGAAGAAAACAGATATAGTGTGTTTGAATCTTTTCTTTCTGTAAGTTCT 1050
 Db 145 ACTGCTTGAAGAAAACAGATATAGTGTGTTTGAATCTTTTCTTTCTGTAAGTTCT 86
 Qy 1051 GTCTGTGCGAAGGCGAGTTGATCAGAGAGCTCAGAGAGGCTTCTGTTTCTAAGTG 1110
 Db 85 GTCTGTGCGAAGGCGAGTTGATCAGAGAGCTCAGAGAGGCTTCTGTTTCTAAGTG 26
 Qy 1111 CCTGAGGGGCGACTCTCTACTGTT 1135
 Db 25 CCTGAGGGGCGACTCTCTACTGTT 1

RESULT 11
 US-09-910-689-71/c
 ; Sequence 71, Application US/09910689
 ; Patent No. US20020081609A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Dillon, Davin C.

APPLICANT: Day, Craig H.
 APPLICANT: Jiang, Yugu
 APPLICANT: Houghton, Raymond L.
 APPLICANT: Mitcham, Jennifer
 APPLICANT: Wang, Tongtong
 APPLICANT: McNeill, Patricia D.
 APPLICANT: Harlocker, Susan L.
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 FILE REFERENCE: 210121.491C6
 CURRENT APPLICATION NUMBER: US/09/910,689
 CURRENT FILING DATE: 2001-07-20
 NUMBER OF SEQ ID NOS: 307
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 71
 LENGTH: 385
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: misc.feature
 LOCATION: 229, 292, 382
 OTHER INFORMATION: n = A,T,C or G
 US-09-910-689-71

Query Match 15.7%; Score 365.2; DB 10; Length 385;
 Best Local Similarity 98.2%; Pred. No. 1.7e-91;
 Matches 378; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

Qy 752 AGAAGCATTCCTCAATCATCTCAGGAGAGAACTTTTACTGTGAGGTGCTGCTAC 811
 Db 385 AGAGCGATTCCTCAATCATCTCAGGAGAGAACTTTTACTGTGAGGTGCTGCTAC 326
 Qy 812 GACTTCTTCGAGGTGCGAGCGAGAT-CGGGGTGGCAGAAATCCAGTTTATGTTCTC 870
 Db 325 GACTTCTTCGAGGTGCGAGCGAGATCCGGGGTGGCAGAAATCCAGTTTATGTTCTC 266
 Qy 871 AGAAGCATTCCTCAATCATCTCAGGAGAGAACTTTTACTGTGAGGTGCTGCTAC 930
 Db 265 AGAAGCATTCCTCAATCATCTCAGGAGAGAACTTTTACTGTGAGGTGCTGCTAC 206
 Qy 931 CACCGGAAATGACTTGGGCAATCATCTGAGTTTGTGATGCGACAGGATTTGGG 990
 Db 205 CACCGGAAATGACTTGGGCAATCATCTGAGTTTGTGATGCGACAGGATTTGGG 146
 Qy 991 ACTGCTTGAAGAAAACAGATATAGTGTGTTTGAATCTTTTCTTTCTGTAAGTTCT 1050
 Db 145 ACTGCTTGAAGAAAACAGATATAGTGTGTTTGAATCTTTTCTTTCTGTAAGTTCT 86
 Qy 1051 GTCTGTGCGAAGGCGAGTTGATCAGAGAGCTCAGAGAGGCTTCTGTTTCTAAGTG 1110
 Db 85 GTCTGTGCGAAGGCGAGTTGATCAGAGAGCTCAGAGAGGCTTCTGTTTCTAAGTG 26
 Qy 1111 CCTGAGGGGCGACTCTCTACTGTT 1135
 Db 25 CCTGAGGGGCGACTCTCTACTGTT 1

RESULT 12
 US-10-010-742-71/c
 ; Sequence 71, Application US/10010742
 ; Patent No. US20020146727A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Dillon, Davin C.
 ; APPLICANT: Day, Craig H.
 ; APPLICANT: Jiang, Yugu
 ; APPLICANT: Houghton, Raymond L.
 ; APPLICANT: Mitcham, Jennifer L.
 ; APPLICANT: Wang, Tongtong
 ; APPLICANT: McNeill, Patricia D.
 ; APPLICANT: Harlocker, Susan L.
 ; APPLICANT: Bennington, Angela Ann
 ; APPLICANT: Zehentner, Barbara
 ; APPLICANT: Fanger, Gary R.

APPLICANT: Retter, Marc W.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.491C7
CURRENT APPLICATION NUMBER: US/10/010,742
CURRENT FILING DATE: 2001-11-30
NUMBER OF SEQ ID NOS: 307
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 71
LENGTH: 385
TYPE: DNA
ORGANISM: Homo sapiens
NAME/KEY: misc feature
LOCATION: 229, 252, 382
OTHER INFORMATION: n = A,T,C or G
US-10-010-742-71

Query Match 15.7%; Score 365.2; DB 12; Length 385;
Best Local Similarity 98.2%; Pred. No. 1.7e-91;
Matches 378; Conservative 0; Mismatches 6; Indels 1; Gaps 1;
QY 752 AGGACGATCCAAATCATCTACGGGAGGAATCTTTTACTGTGGAGGTGGCTGTCTAC 811
DB 385 AGGCGCATCCAAATCATCTACGGGAGGAATCTTTTACTGTGGAGGTGGCTGTCTAC 326
QY 812 GACTCTTCGGAGTGGCAGCCGAGAT- CGGGGTGGCAGAAATCCCAAGTTCATGTGCTC 870
DB 325 GACTCTTCGGAGTGGCAGCCGAGAT- CGGGGTGGCAGAAATCCCAAGTTCATGTGCTC 266
QY 871 AGAAGAGATCAAGCCGCTGTCCTTGTCTAATGCTGCACACCACTTACTGTCTATGG 930
DB 265 AGAAGAGATCAAGCCGCTGTCCTTGTCTAATGCTGCACACCACTTACTGTCTATGG 206
QY 931 CACCGGGAATGACTTGGGCAATCACTGAGTTTGTGTGATCGCAAGGACATTTGGG 990
DB 205 CACCGGGAATGACTTGGGCAATCACTGAGTTTGTGTGATCGCAAGGACATTTGGG 146
QY 991 ACTGCTTGAGAAACAGATAATGATAGTGTGTGTGACTTGTCTTCTGGTAGTTCT 1050
DB 145 ACTGCTTGAGAAACAGATAATGATAGTGTGTGTGACTTGTCTTCTGGTAGTTCT 86
QY 1051 GTCTGCCAAGGCGAGTTGATCAGTGAGTCTCAGGAGAGAGCTTCTGTCTTAAGTGG 1110
DB 85 GTCTGCCAAGGCGAGTTGATCAGTGAGTCTCAGGAGAGAGCTTCTGTCTTAAGTGG 26
QY 1111 CTGTCAGGGGCCACTCTCTACTGTT 1135
DB 25 CTGTCAGGGGCCACTCTCTACTGTT 1

RESULT 13
US-09-745-288-83/c
Sequence 83, Application US/09745288
Patent No. US20010018058A1
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Xu, Jiangchun C.
APPLICANT: Dillon, Davin C.
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
FILE REFERENCE: 210121.446D1
CURRENT APPLICATION NUMBER: US/09/745,288
CURRENT FILING DATE: 2000-12-19
NUMBER OF SEQ ID NOS: 101
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 83
LENGTH: 251
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(251)

OTHER INFORMATION: n = A,T,C or G
US-09-745-288-83

Query Match 10.6%; Score 246.8; DB 10; Length 251;
Best Local Similarity 98.8%; Pred. No. 8.8e-59;
Matches 248; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2081 TTGGGAAGATGTTTATTCTTTTGAGAACCAATATGTTTGCATTAAATGCTTTGATTGT 2140
DB 251 TTGGGAATGTTTATTCTTTTGAGAACCAATATGTTTGCATTAAATGCTTTGATTGT 192
QY 2141 TCATATCAAGAAATGATTGAACGTTCTCAAAACCCCTGTTTACGGTACTTGTGAAGAGGA 2200
DB 191 TCATATCAAGAAATGATTGAACGTTCTCAAAACCCCTGTTTACGGTACTTGTGAAGAGGA 132
QY 2201 GCGGTTTGGGAGAGACCAATTCATCGCTGTCCAAGTGTCTTGTGAAGTGTCTTTTAA 2260
DB 131 GCGGTTTGGGAGAGACCAATTCATCGCTGTCCAAGTGTCTTGTGAAGTGTCTTTTAA 72
QY 2261 CTGGAGAGCTAACCTCAAAATACTTTTTTAACTGCAATCTTATATAATGCGCACACT 2320
DB 71 CTGGAGAGCTAACCTCAAAATATTTTTTAACTGCAATCTTATATAATGCGCACACT 12
QY 2321 ATGCTCCTTAC 2331
DB 11 ATGCTCCTTAC 1

RESULT 14
US-09-960-352-6648
Sequence 6648, Application US/09960352
Patent No. US20020137139A1
GENERAL INFORMATION:
APPLICANT: Warren, Wesley C.
APPLICANT: Tao, Nengbing
APPLICANT: Byatt, John C.
APPLICANT: Mathialagan, Nagappan
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
FILE REFERENCE: 16511.006/37-21(10298)C
CURRENT APPLICATION NUMBER: US/09/960,352
CURRENT FILING DATE: 2001-09-24
NUMBER OF SEQ ID NOS: 15112
SEQ ID NO 6648
LENGTH: 473
TYPE: DNA
ORGANISM: Bos taurus
FEATURE:
NAME/KEY: unsure
LOCATION: (386)
OTHER INFORMATION: unsure at all n locations
OTHER INFORMATION: Clone ID: 29-BOVMS1-009-Q1-E1-H9
US-09-960-352-6648

Query Match 9.9%; Score 230.2; DB 10; Length 473;
Best Local Similarity 78.7%; Pred. No. 5.2e-54;
Matches 354; Conservative 0; Mismatches 79; Indels 17; Gaps 6;
QY 1426 CACTCAGCATCCCAACCGTGCACAGTGAATCTTGTGCTTTGCAAAAGAAAGTGTATCTG 1485
DB 24 CAATCCATCTCTCCGATGCAACCACTTGTGCTTTGCAAAAGAAAGTGTATCTG 83
QY 1486 AAATTCCTCTGTAGAAATTTAGCTTTATACAAATTCAGAGAAATAGCAGTTTCTACTGCCAACTT 1545
DB 84 AAATTCCTGTGTAGCAATTTTCGTTTATATAAATTCAGAAATAGCAGTTTCTACTGCCAACTT 143
QY 1546 TAGTGGGTGAGAAATTTAGTTAGGTGTTTGGATCGG-----ACCTCAGTTTCTGTT 1600
DB 144 ATAGTGGGTGAGAAATTTAGTTAGGTGTTTAGATGTTTATAGATCAGACAATACGTGGGTTTCATT 203
QY 1601 GTTTCCTTTATGTGGTGTCTTAT-ACATGAATCATAGCCAAAACCTTTT---TGGAA 1656
DB 204 CTTTCTTGACGTGTGTAGTTTATATATAAATCATAGCCAAAACCTTTTTCGGGGGA 263

QY 1657 ACTGTGGTGAAGATAGTTGGTCTTTTA-CCCCAGAAAGATCAAGATACCTTGTA 1715
 Db 264 ATAGTCAGTTGAGATCATTAATTTTTTAACCCCAATAAGATCAATTAACCTTGTA 323
 QY 1716 ATAAAGCTGATGACATATATTCATACCTGTTGACCTTGGGT-GAAAGTATGCGAGT 1773
 Db 324 ATAAAGCCGATGATATATTCATACACCTGTTGACCTTGGGTGACATATATGCGCAT 383
 QY 1774 GCGAGACTA-----AGATGATTAACCTACCTGTAATCATATGTTTGAAGAAAGCTGT 1828
 Db 384 GGNAGACTAGTGTGAGATGTTAACTCCCGTGATGATGATGTTGTAAGAAAGCGCTT 443
 QY 1829 TCCCATGCTACAGAGACTTGAAATCAAG 1858
 Db 444 TCAGATGTTGATGGAGCTTGACTTCAG 473

RESULT 15

US-09-954-456-497
 / Sequence 497, Application US/09954456
 / Patent No. US20020115057A1
 / GENERAL INFORMATION:
 / APPLICANT: Young, Paul
 / TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
 / TITLE OF INVENTION: Sets
 / FILE REFERENCE: 689290-76
 / CURRENT APPLICATION NUMBER: US/09/954,456
 / CURRENT FILING DATE: 2001-09-18
 / PRIOR APPLICATION NUMBER: US/60/233,617
 / PRIOR FILING DATE: 2000-09-18
 / PRIOR APPLICATION NUMBER: US/60/234,052
 / PRIOR FILING DATE: 2000-09-20
 / PRIOR APPLICATION NUMBER: US/60/234,923
 / PRIOR FILING DATE: 2000-09-25
 / PRIOR APPLICATION NUMBER: US/60/235,134
 / PRIOR FILING DATE: 2000-09-25
 / PRIOR APPLICATION NUMBER: US/60/235,637
 / PRIOR FILING DATE: 2000-09-26
 / PRIOR APPLICATION NUMBER: US/60/235,638
 / PRIOR FILING DATE: 2000-09-26
 / PRIOR APPLICATION NUMBER: US/60/235,711
 / PRIOR FILING DATE: 2000-09-27
 / PRIOR APPLICATION NUMBER: US/60/235,720
 / PRIOR FILING DATE: 2000-09-27
 / PRIOR APPLICATION NUMBER: US/60/235,840
 / PRIOR FILING DATE: 2000-09-27
 / PRIOR APPLICATION NUMBER: US/60/235,863
 / PRIOR FILING DATE: 2000-09-27
 / NUMBER OF SEQ ID NOS: 2276
 / SOFTWARE: PatentIn version 3.0
 / SEQ ID NO 497
 / LENGTH: 3184
 / TYPE: DNA
 / ORGANISM: Homo sapiens
 US-09-954-456-497

Query Match 8.3%; Score 194.4; DB 10; Length 3184;
 Best Local Similarity 59.9%; Pctd. No. 1,4e-43;
 Matches 374; Conservative 0; Mismatches 226; Indels 24; Gaps 2;

QY 132 TAAGAAAGCAAGATGATCTTTAGAACTTAACATACAGTTTGTGCTCCCTGATGCTGT 191
 Db 178 TCAGAAAGGGGAATGCCAGCCCTAGCATGATGATGATTTCCACTGCTGTTGCTGT 237
 QY 192 GTGGCAAAAGATGATATCTTCAAGGAAAGTGAACAGGCGCAATTTGAGTCCCTCTT 251
 Db 238 GTGGTGAATGTCAGGCTCTTACCAATCAGGAGGTTAAGGAAAAATTTGGGGAGCTGTTT 297
 QY 252 AGGACGATATGACAAAGACATCACTTTCAGATATTTAAGAGCTTCAAGAGAGTCAATA 311
 Db 298 CGGACTTATGATGACTGTGTGACGTTCCAGCTATTTAAGAGTTTCAAGACGTGTCGTA 357

QY 312 AACTTCAGCAACCCCTTCTCCGAGAGATGCCAGGCTCCAGCTGATTAAGACTGATT 371
 Db 358 AACTTCAGCAATCTTAAATCTGCAGCCCGAGCTAGGATAGCTTCATGAAACCAATTC 417
 QY 372 CTGGGAAAGAAATGAAGTTATATTTTGTCTAG-----ACCTTACATATGGA 419
 Db 418 AGAGGGAAAAATTAAGCTCTACTTTGACACAGATTTCAGCTCCAGAGACATGAGAC 477
 QY 420 AGCTCAGACCTGGCTCCGCCAAATCCAGACAGAGTTTGTGATCTTCCCTCCGCTCT 479
 Db 478 AAATGCACTTGGCTCCAGCCCGCTGCCAAGATTTCTCATCTGCGCCCTCTCC 537
 QY 480 CCGCCAGTGGGATGGAACAAAGTGAAGATGCCAGCCCACTATAATCTATGATCTTTA 539
 Db 538 CCACCTGTTAGCTGGCAGCCCATCAAGATGCCAGCCAGTCTTCAACTATGACCTTCTC 597
 QY 540 TATGCAATCTTCAAGCTGGGGCCAGGGGAAAGTAAATTTGACGCGAGGACTGACAC 599
 Db 598 TATGCTGTGGCAAACTAGAGACAGAGAGAGTATAGCTCCATGCAAGGAGTGAAGTCC 657
 QY 600 ACTCCAGCGTGTGTCTCATGATGATGAGTATCAAGAGAGAGAGAGAGAGAGAGAG 659
 Db 658 ACCCAAGTGTCTGTCTGACGTGTGGACAGATGACATAGAG-----GAAAGAA 705
 QY 660 ATGAAAGATGAGAGAGACTTAAGCCAAATTAATCCAGACCCAGAGAGCGGAGTACAG 719
 Db 706 GAGGACCCAAAGACTTCCCAAGCCAAATATCATCCAACTCGCGTCTGCGCTGCCA 765
 QY 720 CCGATCCACCTCAGCTGAACCTGGC 743
 Db 766 CCTCTCGTGTCCAACTGAGCTGCC 789

Search completed: December 14, 2002, 20:01:53
 Job time : 90.3082 secs

241 AGTCCCTCTTTAGAGATGATGACAGAGATCATCCCTTCAGTATTTTAAAGACTTCAAA 300
 Db ATCCCTCTTTTGGAGAGATGACAGAGATCATCTTTCAGTATTTTAAAGACTTCAAA 307
 248 AGTCCCTCTTTTGGAGAGATGACAGAGATCATCTTTCAGTATTTTAAAGACTTCAAA 307
 301 GAGTCAGATTAACCTTACAGCAACCCCTTCTCCGAGCAGATCCAGGCTCCAGTGCATA 360
 Db GAGTCAGATTAACCTTACAGCAACCCCTTCTCCGAGCAGATCCAGGCTCCAGTGCATA 367
 308 GAGTCAGATTAACCTTACAGCAACCCCTTCTCCGAGCAGATCCAGGCTCCAGTGCATA 367
 361 AGACTGAGTTTCTGGGAAAGAAATGATATATTTTCTCAGACCTTACATAGGAA 420
 Db AGACTGAGTTTCTGGGAAAGAAATGATATATTTTCTCAGACCTTACATAGGAA 427
 368 AGACTGAGTTTCTGGGAAAGAAATGATATATTTTCTCAGACCTTACATAGGAA 427
 421 GCTCAACCT -GGCTCCGCCAATCCAGACAGCAGTTCGATCTCCCTCCGCTCT 479
 Db GCTCAACCTCCGCTCCGCCAATCCAGACAGCAGTTCGATCTCCCTCCGCTCT 487
 428 GCTCAACCTCCGCTCCGCCAATCCAGACAGCAGTTCGATCTCCCTCCGCTCT 487
 480 CCSCCAGGAGATGAAACAAAGTGAAGATGCGACCCAGTCATAAATATGATCTCTTA 539
 Db CCSCCAGGAGATGAAACAAAGTGAAGATGCGACCCAGTCATAAATATGATCTCTTA 547
 488 CCSCCAGGAGATGAAACAAAGTGAAGATGCGACCCAGTCATAAATATGATCTCTTA 547
 540 TATGCCATCTCAAGCTGGGGCCAGGGGAAAGATGAAATGACCGACGAGTGCACCC 599
 Db TATGCCATCTCAAGCTGGGGCCAGGGGAAAGATGAAATGACCGACGAGTGCACCC 607
 548 TATGCCATCTCAAGCTGGGGCCAGGGGAAAGATGAAATGACCGACGAGTGCACCC 607
 600 ACTCCAGGAGTGGTGGTCCATGATGAGAGATGATCAAGAGAGAGAGAGAGAGAA 659
 Db ACTCCAGGAGTGGTGGTCCATGATGAGAGATGATCAAGAGAGAGAGAGAGAGAA 667
 608 ACTCCAGGAGTGGTGGTCCATGATGAGAGATGATCAAGAGAGAGAGAGAGAGAA 667
 660 ATGAAAGAAATGAGAGAGATTAAGCCAAAATTTATCCAGCCAGAGGCGGAGTACAG 719
 Db ATGAAAGAAATGAGAGAGATTAAGCCAAAATTTATCCAGCCAGAGGCGGAGTACAG 727
 668 ATGAAAGAAATGAGAGAGATTAAGCCAAAATTTATCCAGCCAGAGGCGGAGTACAG 727
 720 CCGATCCACTCAGCTGAATCTGCGACGCGAGAGAGAGATTCCTCAATCATCTACGCG 729
 Db CCGATCCACTCAGCTGAATCTGCGACGCGAGAGAGAGATTCCTCAATCATCTACGCG 787
 728 CCGATCCACTCAGCTGAATCTGCGACGCGAGAGAGAGATTCCTCAATCATCTACGCG 787
 780 AGGAATCTTTTACTGCTGAGAGTGGCTGTGACGACTTCTTGGAGGTGGAGCCGAGATC 839
 Db AGGAATCTTTTACTGCTGAGAGTGGCTGTGACGACTTCTTGGAGGTGGAGCCGAGATC 847
 788 AGGAATCTTTTACTGCTGAGAGTGGCTGTGACGACTTCTTGGAGGTGGAGCCGAGATC 847
 840 GGGGTGACAGAAATCCAGTTCATGTTGCTCAGAAAGAAATCAAGGCGCTGCTCCCTGT 899
 Db GGGGTGACAGAAATCCAGTTCATGTTGCTCAGAAAGAAATCAAGGCGCTGCTCCCTGT 907
 848 GGGGTGACAGAAATCCAGTTCATGTTGCTCAGAAAGAAATCAAGGCGCTGCTCCCTGT 907
 900 TCTAATCTGACACCACTTACGTTCATGTCGACCCGGGAAATGACTTGGGCAATCATG 959
 Db TCTAATCTGACACCACTTACGTTCATGTCGACCCGGGAAATGACTTGGGCAATCATG 967
 908 TCTAATCTGACACCACTTACGTTCATGTCGACCCGGGAAATGACTTGGGCAATCATG 967
 960 AGTTGTGAGTGCACACAGAGATTTGGAGCTGTCTTGAGAAACAGATATGATAGT 1019
 Db AGTTGTGAGTGCACACAGAGATTTGGAGCTGTCTTGAGAAACAGATATGATAGT 1027
 968 AGTTGTGAGTGCACACAGAGATTTGGAGCTGTCTTGAGAAACAGATATGATAGT 1027
 1020 GTTTTGTACTGTTCTTTCTTGTGAGTTCGTCTGTGCCAAGGGGAGGTTGATCAGTGA 1079
 Db GTTTTGTACTGTTCTTTCTTGTGAGTTCGTCTGTGCCAAGGGGAGGTTGATCAGTGA 1087
 1028 GTTTTGTACTGTTCTTTCTTGTGAGTTCGTCTGTGCCAAGGGGAGGTTGATCAGTGA 1087
 1080 GCTCAGGAGAGAGCTTCTGTTTCTTAAGTGGCTGACAGGGGCACTCTTAATGTAAGA 1139
 Db GCTCAGGAGAGAGCTTCTGTTTCTTAAGTGGCTGACAGGGGCACTCTTAATGTAAGA 1147
 1088 GCTCAGGAGAGAGCTTCTGTTTCTTAAGTGGCTGACAGGGGCACTCTTAATGTAAGA 1147
 1140 AGAGTACACAG 1199
 Db AGAGTACACAG 1207
 1148 AGAGTACACAG 1207
 1200 TGGAAATGTAAGCTTTCTTCTTCTTCTCAGTTCATGTTTGTGATATGATATTAAT 1259
 Db TGGAAATGTAAGCTTTCTTCTTCTTCTCAGTTCATGTTTGTGATATGATATTAAT 1267
 1208 TGGAAATGTAAGCTTTCTTCTTCTTCTCAGTTCATGTTTGTGATATGATATTAAT 1267
 1260 GATTACAGAGATTAACCTTTGTTTCTATATTAAGTACACCGTGTGTTTAACTTTT 1319
 Db GATTACAGAGATTAACCTTTGTTTCTATATTAAGTACACCGTGTGTTTAACTTTT 1327
 1268 GATTACAGAGATTAACCTTTGTTTCTATATTAAGTACACCGTGTGTTTAACTTTT 1327
 1320 TGGAGAGCCAGAGAAAGCTTTGGAAACGATATACCTTTCCAGATTTCCGATTTCTGA 1379

1328 TGGAGAGCCAGAGAAAGCTTTGGAAACGATATACCTTTCCAGATTCGAGTTCTCGA 1387
 1380 CTCTTTCAGACAGACTTGTGCGGAACTCTTCTGGAATGATTCACATCAGATCCCC 1439
 Db CTCTTTCAGACAGACTTGTGCGGAACTCTTCTGGAATGATTCACATCAGATCCCC 1447
 1388 CTCTTTCAGACAGACTTGTGCGGAACTCTTCTGGAATGATTCACATCAGATCCCC 1447
 1440 AACCGTCAACGTTAACTTGTGCTTTTGGCAAAAGAGTGTATGAAATTCCTCTGTAG 1499
 Db AACCGTCAACGTTAACTTGTGCTTTTGGCAAAAGAGTGTATGAAATTCCTCTGTAG 1507
 1448 AACCGTCAACGTTAACTTGTGCTTTTGGCAAAAGAGTGTATGAAATTCCTCTGTAG 1507
 1500 AATTAGCTTATACAAATTCAGAGAAATGACAGTTCACATGCACTTTTATGGGTGAGAA 1559
 Db AATTAGCTTATACAAATTCAGAGAAATGACAGTTCACATGCACTTTTATGGGTGAGAA 1567
 1508 AATTAGCTTATACAAATTCAGAGAAATGACAGTTCACATGCACTTTTATGGGTGAGAA 1567
 1560 AATTAGCTTATGTTTGGATGCGACCTCAGTTCGTGTGTTCTTTATGATGATGAT 1619
 Db AATTAGCTTATGTTTGGATGCGACCTCAGTTCGTGTGTTCTTTATGATGATGAT 1627
 1568 AATTAGCTTATGTTTGGATGCGACCTCAGTTCGTGTGTTCTTTATGATGATGAT 1627
 1620 TTTATACATGATATCATAGCCAAAACCTTTTGGAAAAGTGTGAGATGATGATGAT 1679
 Db TTTATACATGATATCATAGCCAAAACCTTTTGGAAAAGTGTGAGATGATGATGAT 1687
 1628 TTTATACATGATATCATAGCCAAAACCTTTTGGAAAAGTGTGAGATGATGATGAT 1687
 1680 CTTTAAACCCAGAGAGATCAAGATACACTTGAATTAAGAGATGATGATGATGAT 1729
 Db CTTTAAACCCAGAGAGATCAAGATACACTTGAATTAAGAGATGATGATGATGAT 1747
 1688 CTTTAAACCCAGAGAGATCAAGATACACTTGAATTAAGAGATGATGATGATGAT 1747
 1740 AACTTGTACACTTGGTGGAAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1799
 Db AACTTGTACACTTGGTGGAAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1807
 1748 AACTTGTACACTTGGTGGAAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1807
 1800 TGTGAATCATATGTTGTGAGAAAGCTTTCATGTTTAAACAGAGCTTGAATTCAAAGC 1859
 Db TGTGAATCATATGTTGTGAGAAAGCTTTCATGTTTAAACAGAGCTTGAATTCAAAGC 1867
 1808 TGTGAATCATATGTTGTGAGAAAGCTTTCATGTTTAAACAGAGCTTGAATTCAAAGC 1867
 1860 ATGTCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1919
 Db ATGTCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1927
 1868 ATGTCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1927
 1920 ATTCCTGATGAAATGTTATATATATATATATATATATATATATATATATATATAT 1979
 Db ATTCCTGATGAAATGTTATATATATATATATATATATATATATATATATATATAT 1987
 1928 ATTCCTGATGAAATGTTATATATATATATATATATATATATATATATATATATAT 1987
 1980 TATGTCGTGATTAATGAT 2039
 Db TATGTCGTGATTAATGAT 2047
 1988 TATGTCGTGATTAATGAT 2047
 2040 TTTGAATATCTTAATGCAAGTAAATGATGATGATGATGATGATGATGATGATGATGAT 2099
 Db TTTGAATATCTTAATGCAAGTAAATGATGATGATGATGATGATGATGATGATGATGAT 2107
 2048 TTTGAATATCTTAATGCAAGTAAATGATGATGATGATGATGATGATGATGATGATGAT 2107
 2100 CTTTGAAGAGAAATATGTTTGCATTAATGCTTTGATGTTTGTTCATATCAAGAAATGAT 2159
 Db CTTTGAAGAGAAATATGTTTGCATTAATGCTTTGATGTTTGTTCATATCAAGAAATGAT 2167
 2108 CTTTGAAGAGAAATATGTTTGCATTAATGCTTTGATGTTTGTTCATATCAAGAAATGAT 2167
 2160 GAACGTTTCAAAACCTGTTTAAGGACTTGTGATGAGAGAGAGAGAGAGAGAGAGAGAG 2219
 Db GAACGTTTCAAAACCTGTTTAAGGACTTGTGATGAGAGAGAGAGAGAGAGAGAGAGAG 2227
 2168 GAACGTTTCAAAACCTGTTTAAGGACTTGTGATGAGAGAGAGAGAGAGAGAGAGAGAG 2227
 2220 TTTGATGCTGTCAGAGTGTCTTGTGATGATGATGATGATGATGATGATGATGATGAT 2279
 Db TTTGATGCTGTCAGAGTGTCTTGTGATGATGATGATGATGATGATGATGATGATGAT 2287
 2228 TTTGATGCTGTCAGAGTGTCTTGTGATGATGATGATGATGATGATGATGATGATGAT 2287
 2280 AATATCTTTTAACTGATCTTAATTAATGAGGACAGATGATGATGATGATGATGATGAT 2331
 Db AATATCTTTTAACTGATCTTAATTAATGAGGACAGATGATGATGATGATGATGATGAT 2339

RESULT 2
 US-09-724-676-12747
 ; Sequence 12747, Application US/09724676
 ; GENERAL INFORMATION:
 ; APPLICANT: Compugen LTD

;
; TITLE UP INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12747
; LENGTH: 2297
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (11)..(11)
; OTHER INFORMATION: n is a,c,g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (18)..(18)
; OTHER INFORMATION: n is a,c,g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (23)..(23)
; OTHER INFORMATION: n is a,c,g, or t
; US-09-724-676-12747

Query Match 95.0%; Score 2214.8; DB 5; Length 2297;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2227; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 102 TGACTCGTGGGTCTGTAGCGCTTTTCACTGTAAAGAAAGCAAGATGATTTTAAAGAACTTT 161
DB 21 TANCTCGTGGGTCTGTAGCGCTTTTCACTGTAAAGAAAGCAAGATGATTTTAAAGAACTTT 80

QY 162 AACTACAGTTTACGTTCCCTGATGCTGTGGGCAAAACAGTGATATCTTCAGCGAAAGT 221
DB 81 AACTACAGTTTACGTTCCCTGATGCTGTGGGCAAAACAGTGATATCTTCAGCGAAAGT 140

QY 222 GAAACAGCGGCGCAAAATTTGAGTCCCTCTTTAGGACGTATGCAAGACATCACTTTTTCAG 281
DB 141 GAAACAGCGGCGCAAAATTTGAGTCCCTCTTTAGGACGTATGCAAGACATCACTTTTTCAG 200

QY 282 TATTTTAAAGGCTTCAAAACAGTCAAGATAAATCTTCAGCAACCCCTTCTCCGACGAGAT 341
DB 201 TATTTTAAAGGCTTCAAAACAGTCAAGATAAATCTTCAGCAACCCCTTCTCCGACGAGAT 260

QY 342 GCAGGCTCCAGTGCATAGAGTGTCTGGGAAAGGAAATGAAGTTATTTTGTCT 401
DB 261 GCAGGCTCCAGTGCATAGAGTGTCTGGGAAAGGAAATGAAGTTATTTTGTCT 320

QY 402 CAGACCTTACATAGGAAGCTCACACCTGGCTCCGCAAAATCCAGCAAGCAGTTTCTG 461
DB 321 CAGACCTTACATAGGAAGCTCACACCTGGCTCCGCAAAATCCAGCAAGCAGTTTCTG 380

QY 462 ATCTCCCTCCGCTCTCCGCAAGTGGGATGGAACAAAGTGGAGATGCGACCCAGTC 521
DB 381 ATCTCCCTCCGCTCTCCGCAAGTGGGATGGAACAAAGTGGAGATGCGACCCAGTC 440

QY 522 ATAACTAGATCTCTTATATGCCATCTCAAGCTGGGGCCAGGGGAAAGATGATGAATTG 581
DB 441 ATAACTAGATCTCTTATATGCCATCTCAAGCTGGGGCCAGGGGAAAGATGATGAATTG 500

QY 582 CACGAGGAGCTGACACCACTCCAGCGTGGTGGTCCATGTATGTGAGAGTCAAGAG 641
DB 501 CACGAGGAGCTGACACCACTCCAGCGTGGTGGTCCATGTATGTGAGAGTCAAGAG 560

QY 642 AAGGAGGAAGAGAGAAATGGAAGATGAGGAGACCTTAAGCCAAATTTATCCAGACC 701
DB 561 AAGGAGGAAGAGAGAAATGGAAGATGAGGAGACCTTAAGCCAAATTTATCCAGACC 620

QY 702 AGGAGCGGAGTACACCGGATCCACCTCAGCTGAACTGGCACCGGACGAGCGCAT 761
DB 621 AGGAGCGGAGTACACCGGATCCACCTCAGCTGAACTGGCACCGGACGAGCGCAT 680

QY 762 CCAATATCATCTCACGGGGAGGAATCTTTTACTGTGGAGGTGGCTGGTCAACGACTTCTTCG 821

DB 681 CCAATATCATCTCACGGGGAGGAATCTTTTACTGTGGAGGTGGCTGGTCAACGACTTCTTCG 740
QY 822 GAGGTGGCAGCCGAGATCGGGGTGGCAGAAATCCCAAGTTTCAATGTTCTCAGAGAGAAATC 881
DB 741 GAGGTGGCAGCCGAGATCGGGGTGGCAGAAATCCCAAGTTTCAATGTTCTCAGAGAGAAATC 800
QY 882 AAGGCCGTGTCCCTTTCTAATCTGCACACCAAGTTTCTTCTTAAGTGGCCCTGCAGGGGAT 941
DB 801 AAGGCCGTGTCCCTTTCTAATCTGCACACCAAGTTTCTTCTTAAGTGGCCCTGCAGGGGAT 860
QY 942 GACTTGGGCCAATCACTGAGTTTGTGTGATCGCAACAGGACATTTGGGACTGTCTTGAG 1001
DB 861 GACTTGGGCCAATCACTGAGTTTGTGTGATCGCAACAGGACATTTGGGACTGTCTTGAG 920
QY 1002 AAAACAGATATGATAGTGTGTGTGATCTTTTCTTTCTTTCTGTGTAGTGTCTGTGTCGAA 1061
DB 921 AAAACAGATATGATAGTGTGTGTGATCTTTTCTTTCTTTCTGTGTAGTGTCTGTGTCGAA 980
QY 1062 GGGCAGGTTGATCAGTGAGCTCAGGAGAGAGCTTCTCTGTTTCTTAAGTGGCCCTGCAGGGGC 1121
DB 981 GGGCAGGTTGATCAGTGAGCTCAGGAGAGAGCTTCTCTGTTTCTTAAGTGGCCCTGCAGGGGC 1040
QY 1122 CACTCTCTACTGTAGGAAGAGGTACCAAGGAGCCGCTAGTGCAGAGAGGTTGTGAA 1181
DB 1041 CACTCTCTACTGTAGGAAGAGGTACCAAGGAGCCGCTAGTGCAGAGAGGTTGTGAA 1100
QY 1182 AACAGCAGCAATGCAATGTGAAATTTAGGAGTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 1241
DB 1101 AACAGCAGCAATGCAATGTGAAATTTAGGAGTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 1160
QY 1242 TTGTGCAATGATATTAACAGACTTAACAGACTTAACAGACTTAACAGACTTAACAGACTTAACAG 1301
DB 1161 TTGTGCAATGATATTAACAGACTTAACAGACTTAACAGACTTAACAGACTTAACAGACTTAACAG 1220
QY 1302 TTGTGTTTTCATCTTTTGGGAGCCAGGAAAGCGTTTGGGAAACGATATCACTTTTCCC 1361
DB 1221 TTGTGTTTTCATCTTTTGGGAGCCAGGAAAGCGTTTGGGAAACGATATCACTTTTCCC 1280
QY 1362 AGATTCTCGGATCTCGACTCTTTTGGCAACAGCAGCTTGTCTTGGGAACTTTCTTCTTGGAAATG 1421
DB 1281 AGATTCTCGGATCTCGACTCTTTTGGCAACAGCAGCTTGTCTTGGGAACTTTCTTCTTGGAAATG 1340
QY 1422 CATTCACCTCAGCATCCCCAACCCGTCGAAACGTTAACTTTGCTTTTGCAGAAAGAGTTGA 1481
DB 1341 CATTCACCTCAGCATCCCCAACCCGTCGAAACGTTAACTTTGCTTTTGCAGAAAGAGTTGA 1400
QY 1482 TCTGAAATTCCTCTGTGAGATTTAGCTTTATACAAATCAGAGAAATAGCAGTTTCTACGTGCA 1541
DB 1401 TCTGAAATTCCTCTGTGAGATTTAGCTTTATACAAATCAGAGAAATAGCAGTTTCTACGTGCA 1460
QY 1542 ACTTTTGTGGGTGAGAAATTTTGTAGTTTGTGGATCGGACCTCAGTTTCTGTGTTG 1601
DB 1461 ACTTTTGTGGGTGAGAAATTTTGTAGTTTGTGGATCGGACCTCAGTTTCTGTGTTG 1520
QY 1602 TTTCTTTTGTGGGTGAGAAATTTTGTAGTTTGTGGATCGGACCTCAGTTTCTGTGTTG 1660
DB 1521 TTTCTTTTGTGGGTGAGAAATTTTGTAGTTTGTGGATCGGACCTCAGTTTCTGTGTTG 1580
QY 1661 TTGTTTGTAGATAGTTGTTCTTTTACCCACGAGACATCAAGATACACTTTGTAATAAAA 1720
DB 1581 TTGTTTGTAGATAGTTGTTCTTTTACCCACGAGACATCAAGATACACTTTGTAATAAAA 1640
QY 1721 GCTGTATAGCATATATTCATACCTGTTGTCATCTGGGTGAAAAGTATGGCAGTGGGAGAC 1780
DB 1641 GCTGTATAGCATATATTCATACCTGTTGTCATCTGGGTGAAAAGTATGGCAGTGGGAGAC 1700
QY 1781 TAAGATGTATTAACCTACCTGTGAAATCATATGTTGTAGGAAAAGCTGTTCCCATGTCTAA 1840
DB 1701 TAAGATGTATTAACCTACCTGTGAAATCATATGTTGTAGGAAAAGCTGTTCCCATGTCTAA 1760
QY 1841 CAGGACTTGAATTCAGAGCATGTCAAAGTGAATAGATCTGTGGCGATATGAGAGGAT 1900

Db 1761 CAGACTTGAATTCMAAGCATGTCAGATAGATCTGTGGCATATGAGAGGAT 1820
Qy 1901 GAGAGGCTTTCCCTTCATTCCTGATGGAAATGTTACTAGGTTAACTTTGTAAT 1960
Db 1821 GAGAGGCTTTCCCTTCATTCCTGATGGAAATGTTACTAGGTTAACTTTGTAAT 1880
Qy 1961 TTTTCTAGTTGATAGTATGTCTGTAAATAGGTAATATTTTGGCTTACATAC 2020
Db 1881 TTTTCTAGTTGATAGTATGTCTGTAAATAGGTAATATTTTGGCTTACATAC 1940
Qy 2021 CGTAACTGTTTGTCTATTTGTAATCTTAAATGTCATTAATGCTTTGGAAA 2080
Db 1941 CGTAACTGTTTGTCTATTTGTAATCTTAAATGTCATTAATGCTTTGGAAA 2000
Qy 2081 TTTGGAAGATGTTTATCTTTGAGAGCAAAATATGTTGATTAATGCTTTGATGT 2140
Db 2001 TTTGGAAGATGTTTATCTTTGAGAGCAAAATATGTTGATTAATGCTTTGATGT 2060
Qy 2141 TCATATCAAGAAATGATTTGAACGTTCTCAAACTGTTTACGGTACTTGTAAAGAGGA 2200
Db 2061 TCATATCAAGAAATGATTTGAACGTTCTCAAACTGTTTACGGTACTTGTAAAGAGGA 2120
Qy 2201 GCCGGTTTGGAGAGACCATTCATGCTGTCCAAAGTGTCTTTGTTAAAGTCTTTTAA 2260
Db 2121 GCCGGTTTGGAGAGACCATTCATGCTGTCCAAAGTGTCTTTGTTAAAGTCTTTTAA 2180
Qy 2261 CTGAGAGGCTAATCTCAAAATCTTTTAACTGATCTAATTAATGAGGACAGT 2320
Db 2181 CTGAGAGGCTAATCTCAAAATCTTTTAACTGATCTAATTAATGAGGACAGT 2240
Qy 2321 ATGCTCTTAC 2331
Db 2241 ATGCTCTTAC 2251

RESULT 3
US-09-724-676A-12747
; Sequence 12747, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12747
; LENGTH: 2297
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (11)..(11)
; OTHER INFORMATION: n is a,c,g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (18)..(18)
; OTHER INFORMATION: n is a,c,g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (23)..(23)
; OTHER INFORMATION: n is a,c,g, or t
US-09-724-676A-12747

Query Match 95.0%; Score 2214.8; DB 5; Length 2297;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2227; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
Qy 102 TGACGCGTGGGTCTGTAGCGCTTCACTGTAAAGCAAGATGATTTTAAACCTT 161
Db 21 TATGCTGCTGGGTCTGTAGCGCTTCACTGTAAAGCAAGATGATTTTAAACCTT 80
Qy 162 AACTACAGTTTATAGCTCCCTGATTCCTGTGTGGCAAAAGTATATCTTACAGGAAAGT 221

Db 81 AACTACAGTTTATAGCTCCCTGATTCCTGTGTGGCAAAAGTATATCTTACAGGAAAGT 140
Qy 222 GAAACCAAGGCGCAATTTGAGTCCCTCTTTAGACGTATGACAAGACATCACCTTTCA 281
Db 141 GAAACCAAGGCGCAATTTGAGTCCCTCTTTAGACGTATGACAAGACATCACCTTTCA 200
Qy 282 TATTTTAAAGCTTCAACGATCAGAAATTAACCTTACGCAACCCCTTCTCCGACGAT 341
Db 201 TATTTTAAAGCTTCAACGATCAGAAATTAACCTTACGCAACCCCTTCTCCGACGAT 260
Qy 342 GCCAGGCTCCGAGCTGATTAAGATGAGTTCTGGGAAAGAAATGAAATTAATTTTGT 401
Db 261 GCCAGGCTCCGAGCTGATTAAGATGAGTTCTGGGAAAGAAATGAAATTAATTTTGT 320
Qy 402 CAGACCTTACACATAGAAAGCTCACACCTGCTCCGCAATCCAGACAAGACGTTTCTG 461
Db 321 CAGACCTTACACATAGAAAGCTCACACCTGCTCCGCAATCCAGACAAGACGTTTCTG 380
Qy 462 ATCTCCCTCCGCTCTCCGCTCAGTGGGATGAAACAAGTGAAGATGCCACCCAGTC 521
Db 381 ATCTCCCTCCGCTCTCCGCTCAGTGGGATGAAACAAGTGAAGATGCCACCCAGTC 440
Qy 522 ATTAACATAGATCTCTTATATGCAATCCAAAGCTGGGCGCAGGAGAAAGTATGAATTTG 581
Db 441 ATTAACATAGATCTCTTATATGCAATCTCCAAAGCTGGGCGCAGGAGAAAGTATGAATTTG 500
Qy 582 CACGAGGAGCTGACACCACTCCAGCTGGTGTCCATGTATGTAGAGTATCAAGAG 641
Db 501 CACGAGGAGCTGACACCACTCCAGCTGGTGTCCATGTATGTAGAGTATCAAGAG 560
Qy 642 AAGAGGAAGAAAGAAATGAAAGATGAGAGACCTTAAGCCAAAATATATCCAGACC 701
Db 561 AAGAGGAAGAAAGAAATGAAAGATGAGAGACCTTAAGCCAAAATATATCCAGACC 620
Qy 702 AAGAGGCGGAGATACAGCGGATCCACCTCAGCTGAATCTGACGCGACGAGACGATT 761
Db 621 AAGAGGCGGAGATACAGCGGATCCACCTCAGCTGAATCTGACGCGACGAGACGATT 680
Qy 762 CCMAATCATACTACGCGGAGAAATCTTTTACTGTGAGGCTGTGTGTACGACTTTCTG 821
Db 681 CCMAATCATACTACGCGGAGAAATCTTTTACTGTGAGGCTGTGTGTACGACTTTCTG 740
Qy 822 GAGGTGGCAGCCGAGATCGGGTGGCAGAAATCCAGTTCATGTTGCACAGAGAAATC 881
Db 741 GAGGTGGCAGCCGAGATCGGGTGGCAGAAATCCAGTTCATGTTGCACAGAGAAATC 800
Qy 882 AAGGCGGTGCTCCCTGTCTTAATGCTGACACAGATTACTGTTTATGAGCCCGGAAAT 941
Db 801 AAGGCGGTGCTCCCTGTCTTAATGCTGACACAGATTACTGTTTATGAGCCCGGAAAT 860
Qy 942 GACTTGGGCAATCACTGAGTTGTGTGTATGTCACAAAGACATTTGGAGCTGTCTTGA 1001
Db 861 GACTTGGGCAATCACTGAGTTGTGTGTATGTCACAAAGACATTTGGAGCTGTCTTGA 920
Qy 1002 AAAACGATTAATAGTATGTTTGTACTTCTTCTGTAGGTTCTGTGTGCCAA 1061
Db 921 AAAACGATTAATAGTATGTTTGTACTTCTTCTGTAGGTTCTGTGTGCCAA 980
Qy 1062 GGGCAGGTTGATCAGTGAAGCTCAGAGAGAGCTTCTGTTTCTAAGTGCCTGCAAGGGC 1121
Db 981 GGGCAGGTTGATCAGTGAAGCTCAGAGAGAGCTTCTGTTTCTAAGTGCCTGCAAGGGC 1040
Qy 1122 CACTCTCTACTGTAGAAAGAGGTACACAGAGAGCGCTATGTGACAGAGGTGTGAA 1181
Db 1041 CACTCTCTACTGTAGAAAGAGGTACACAGAGAGCGCTATGTGACAGAGGTGTGAA 1100
Qy 1182 AAGAGCAAGATGCAATGTGAAATTTGTAGCGTTCTTCTTCCCTCATGTTCTCATGT 1241
Db 1101 AAGAGCAAGATGCAATGTGAAATTTGTAGCGTTCTTCTTCCCTCATGTTCTCATGT 1160
Qy 1242 TTGTGATGTATATTAAGTATTTTACAAAGCTAACTTTGTGTATATTAAGTTACCGG 1301

Db	1161	TTGTGCATGTATATTA	CTACTGATTTTCAAGAC	TAACTTTGTTTCGTATAT	AAAGTTACACCG	1322
Qy	1302	TTGTTGTTTTACATCT	TTTGGGAAGCAGGAAG	CGGTTTGGAAAACGTAT	CACCTTTCCC	1361
Db	1221	TTGTTGTTTTACATCT	TTTGGGAAGCAGGAAG	CGGTTTGGAAAACGTAT	CACCTTTCCC	1280
Qy	1362	AGATTCTCGGATTC	TCGACCTTTTGCAAC	CAGCACTTCCTTCGCGAA	CTCTTCTCGGAATG	1421
Db	1381	AGATTCTCGGATTC	TCGACCTTTTGCAAC	CAGCACTTCCTTCGCGAA	CTCTTCTCGGAATG	1340
Qy	1422	CATTCACTCAGCATC	CCCCAACCGTGCAAC	CGTGTAACTTTTGCAAAA	GAAGTTGA	1481
Db	1341	CATTCACTCAGCATC	CCCCAACCGTGCAAC	CGTGTAACTTTTGCAAAA	GAAGTTGA	1400
Qy	1482	TCTGAANAATCCT	CTGTAGAAATTTAG	CTATTACAAATTCAGAGA	TAGCAGTTTCACATGCCA	1541
Db	1401	TCTGAANAATCCT	CTGTAGAAATTTAG	CTATTACAAATTCAGAGA	TAGCAGTTTCACATGCCA	1460
Qy	1542	ACTTTAGTCGGTG	GAGAAATTTTAGTT	TAGGTCCTTGGGAT	CGGACCTCAGTTTCTGTG	1601
Db	1461	ACTTTAGTCGGTG	GAGAAATTTTAGTT	TAGGTCCTTGGGAT	CGGACCTCAGTTTCTGTG	1520
Qy	1602	TTTCTTTTATGT	CGTGTTTCTATAC	ATGAATCATAGGCC	AAAACTTTTTCG-AAA	1660
Db	1521	TTTCTTTTATGT	CGTGTTTCTATAC	ATGAATCATAGGCC	AAAACTTTTTCG-AAA	1580
Qy	1661	TTGGTTTGAGATAG	TGGTTCCTTTTAC	CCCCGGAACATCA	AGATACCTTGTAAATAA	1720
Db	1581	TTGGTTTGAGATAG	TGGTTCCTTTTAC	CCCCGGAACATCA	AGATACCTTGTAAATAA	1640
Qy	1721	GCTGATAGCATAT	TATTCATACCTG	TGTACCTTGGTG	GAAAGTATGCGAGTGGAGAC	1780
Db	1641	GCTGATAGCATAT	TATTCATACCTG	TGTACCTTGGTG	GAAAGTATGCGAGTGGAGAC	1700
Qy	1781	TAAGATGTAATA	ACCTACCTGTGA	ATCATATGTTAGG	AAAAAGCTGTTCCCATGCTCAA	1840
Db	1701	TAAGATGTAATA	ACCTACCTGTGA	ATCATATGTTAGG	AAAAAGCTGTTCCCATGCTCAA	1760
Qy	1841	CAGGACTCGAAT	TCAAAGCATG	CAAGTGGATAG	TCTGTGGCGATATGAGAGGAT	1900
Db	1761	CAGGACTCGAAT	TCAAAGCATG	CAAGTGGATAG	TCTGTGGCGATATGAGAGGAT	1820
Qy	1901	GCAGTGCCTTT	CCCCATTCCTG	ATGGAATGTTAT	ACTAGGTAAACATTTGCTAAT	1960
Db	1821	GCAGTGCCTTT	CCCCATTCCTG	ATGGAATGTTAT	ACTAGGTAAACATTTGCTAAT	1880
Qy	1961	TTTTTCTAGTCT	TAATGTATGCTG	TAAATAGGTAAT	TAATTTGCGCTTACAATAC	2020
Db	1881	TTTTTCTAGTCT	TAATGTATGCTG	TAAATAGGTAAT	TAATTTGCGCTTACAATAC	1940
Qy	2021	CGTAACAATGTT	GTGTCATTTGAA	ATACCTTAATG	CGCAATGCTTTTAA	2080
Db	1941	CGTAACAATGTT	GTGTCATTTGAA	ATACCTTAATG	CGCAATGCTTTTAA	2000
Qy	2081	TTTGAACATG	TTTTATCTTTG	AGAGCAAAATAT	GTTCATTAATGCTTTGATGT	2140
Db	2001	TTTGAACATG	TTTTATCTTTG	AGAGCAAAATAT	GTTCATTAATGCTTTGATGT	2060
Qy	2141	TCATATCAAGAA	ATTGATTGAAC	GTCTCAAACCTGTT	TACCGTACTTGGAAGGGA	2200
Db	2061	TCATATCAAGAA	ATTGATTGAAC	GTCTCAAACCTGTT	TACCGTACTTGGAAGGGA	2120
Qy	2201	GCCGGTTGGG	AGAGCAATTCG	ATCGTCCAAAGT	TTTTCTTGTTAAAGTCTTTTAA	2260
Db	2121	GCCGGTTGGG	AGAGCAATTCG	ATCGTCCAAAGT	TTTTCTTGTTAAAGTCTTTTAA	2180
Qy	2261	CTGAGAGGCT	TAACTCAAATAT	CTTTTTTA	CTGCAATCTATATAAATGGGCACAGT	2320
Db	2181	CTGAGAGGCT	TAACTCAAATAT	CTTTTTTA	CTGCAATCTATATAAATGGGCACAGT	2240
Qy	2321	ATGCTCCTTAC	2331			
Db	2241	ATGCTCCTTAC	2251			

RESULT 4

```

US-09-724-676-12752
; Sequence 12752, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: CompuGen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 CompuGen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97232
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12752
; LENGTH: 1943
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676-12752

```

Query Match	78.7%;	Score 1833.4;	DB 5;	Length 1943;	
Best Local Similarity	99.9%;	Pred. No. 0;			
Matches 1845;	Conservative	0;	Mismatches	1;	Indels
					1;
QY 486	GTGGGATGGAAACAAGTGAAGATGCGAGCCCGAGTCATAAACTATGATCTCTTATATGCC	545			
Db					
QY 51	GTGGGATGGAAACAAGTGAAGATGCGAGCCCGAGTCATAAACTATGATCTCTTATATGCC	110			
Db					
QY 546	ATCTCCAAGCTGGGCCAGGGGAAAAGATGTAATTTGACGCGAGCGACTGACACCACTCCC	605			
Db					
QY 111	ATCTCCAAGCTGGGCCAGGGGAAAAGATGTAATTTGACGCGAGCGACTGACACCACTCCC	170			
Db					
QY 606	AGCGTGGTCCATGTATGTGAGAGTGATCAAGAGAAAGGAGGAAGAGAGGAATGGAA	665			
Db					
QY 171	AGCGTGGTGGTCCATGTATGTGAGAGTGATCAAGAGAAAGGAGGAAGAGAGGAATGGAA	230			
Db					
QY 666	AGAAATGAGGAGACCTTAAGCCAAAAATTATCCAGACCGAGGCGCGAGTACACGCCGATC	725			
Db					
QY 231	AGAAATGAGGAGACCTTAAGCCAAAAATTATCCAGACCGAGGCGCGAGTACACGCCGATC	290			
QY 726	CACCTCAGCTGAATCGGCACGCGAGCGAGCGCATTCCTCAATCATCTACCGGAGGAAT	785			
Db					
QY 291	CACCTCAGCTGNACTGGCACGCGAGCGAGCGCATTCCTCAATCATCTACCGGAGGAAT	350			
Db					
QY 786	CTTTTACTGTGGAGGTGGTGGTCAAGCTCTTTCGGAGGTGGCAGCGAGATCGGGGTG	845			
Db					
QY 351	CTTTTACTGTGGAGGTGGTGGTCAAGCTCTTTCGGAGGTGGCAGCGAGATCGGGGTG	410			
Db					
QY 846	GCAGAAATCCCAAGTTCATGTTGCTCAGAGAGAAATCAAGCGCGTGTCCCTTGTTCTAAT	905			
Db					
QY 411	GCAGAAATCCCAAGTTCATGTTGCTCAGAGAGAAATCAAGCGCGTGTCCCTTGTTCTAAT	470			
QY 906	GCTGCACACCAAGTTACTGTTTCATCGCACCCGGGAATACACTTGGGCCAATCACTGAGTTG	965			
Db					
QY 471	GCTGCACACCAAGTTACTGTTTCATCGCACCCGGGAATACACTTGGGCCAATCACTGAGTTG	530			
Db					
QY 966	TGGTGATCGCAACAAGGACATTTGGGACTGTCTTGAGAAAAACAGATATGATAGTGTGTTG	1025			
Db					
QY 531	TGGTGATCGCAACAAGGACATTTGGGACTGTCTTGAGAAAAACAGATATGATAGTGTGTTG	590			
QY 1026	TACTGTGTTCTTTTCTGTTAGTGTGTTCTGTCTGTGCAAGGGCAGGTTGATCAGTGAGCTCAG	1085			
Db					
QY 591	TACTGTGTTCTTTTCTGTTAGTGTGTTCTGTCTGTGCAAGGGCAGGTTGATCAGTGAGCTCAG	650			
Db					
QY 1086	GAGAGAGCTTCTGTTTCTTAAGTGGCTCGCAGGGGCCATCTCTACTGTTAGGAAGAGGT	1145			
Db					
QY 651	GAGAGAGCTTCTGTTTCTTAAGTGGCTCGCAGGGGCCATCTCTACTGTTAGGAAGAGGT	710			
Db					
QY 1146	ACCAAGGAAGCCCGCTAGTGTGACAGAGAGGTTGTGAAAAACAGCAGCAATGCAATGTGGAAA	1205			
Db					
QY 711	ACCAAGGAAGCCCGCTAGTGTGACAGAGAGGTTGTGAAAAACAGCAGCAATGCAATGTGGAAA	770			
Db					
QY 1206	TTGTAGGGTTTCTTTCTTCCCTCATGTTCTCATGTTTGTGATGTPATATTACTGATTTA	1265			
Db					

Db 771 TTGTAGCGTTTCCTTCTTCCCTCATGTTCTCATGTTTGCGCATATATATCTGATTTA 830
 Qy 1266 CAAGACTAACCTTGTGTGTATATTAAGTTACACCGTGTGTGTTTATCATCTTTGGGAA 1325
 Db 831 CAAGACTAACCTTGTGTGTATATTAAGTTACACCGTGTGTGTTTATCATCTTTGGGAA 890
 Qy 1326 GCCAGGAAAGCGTTTGGAAAGATATCATCTTTCCGAGATCTCGGATCTCGACCTCTT 1385
 Db 891 GCCAGGAAAGCGTTTGGAAAGATATCATCTTTCCGAGATCTCGGATCTCGACCTCTT 950
 Qy 1386 GCAGAGACCTGCTTGGGGAAGCTCTCTGGAATGATCATCTGAGATCCCAACCGT 1445
 Db 951 GCAGAGACCTGCTTGGGGAAGCTCTCTGGAATGATCATCTGAGATCCCAACCGT 1010
 Qy 1446 GCAGAGACCTGCTTGTGCTTTTGGAAAGAGTGTGAATTCCTCTGTGAATTTA 1505
 Db 1011 GCAGAGACCTGCTTGTGCTTTTGGAAAGAGTGTGAATTCCTCTGTGAATTTA 1070
 Qy 1506 GCTTATACATTCAGAGATAGACATTTTCACTGCAACCTTTAGTGGGAGAAATTTTA 1565
 Db 1071 GCTTATACATTCAGAGATAGACATTTTCACTGCAACCTTTAGTGGGAGAAATTTTA 1130
 Qy 1566 GTTTAGGTGTTTGGGATCGGACCTGATTTCTGTGTTCTTTTATGTGTGTTCTAT 1625
 Db 1131 GTTTAGGTGTTTGGGATCGGACCTGATTTCTGTGTTCTTTTATGTGTGTTCTAT 1190
 Qy 1626 ACATGATATAGCCAAAAAATTTTGG-AAAAGTGTGTGAATAGTGTGTTCTTT 1684
 Db 1191 ACATGATATAGCCAAAAAATTTTGGAAACGTGTGTGAATAGTGTGTTCTTT 1250
 Qy 1685 ACCCCAGGAAGATATCAAGATATACCTGTAAATTAAGCTGATATATTCATACCTG 1744
 Db 1251 ACCCCAGGAAGATATCAAGATATACCTGTAAATTAAGCTGATATATTCATACCTG 1310
 Qy 1745 TTGTACACTTGGGTGAAAAGTATGCGAGTGGAGACTAAGATGTATTAACCTGTGA 1804
 Db 1311 TTGTACACTTGGGTGAAAAGTATGCGAGTGGAGACTAAGATGTATTAACCTGTGA 1370
 Qy 1805 ATCATATGTGTGAGAAAAGCTGTCCCATGTCTAACAGGACTGAAATTCMAAGCATGTC 1864
 Db 1371 ATCATATGTGTGAGAAAAGCTGTCCCATGTCTAACAGGACTGAAATTCMAAGCATGTC 1430
 Qy 1865 AAGGTGATAGATCTGTGGCATATGAGAGGATGCACTGCTTCCCATTCATCC 1924
 Db 1431 AAGGTGATAGATCTGTGGCATATGAGAGGATGCACTGCTTCCCATTCATCC 1490
 Qy 1925 TGATGAAATGTTATATCTAGGTTAACATTTGTAATTTTTTTCTAGTGTAAATGTATGT 1984
 Db 1491 TGATGAAATGTTATATCTAGGTTAACATTTGTAATTTTTTTCTAGTGTAAATGTATGT 1550
 Qy 1985 CTGTAAATAGTATATATTTTGGCCCTTACAAATACCGTAACAAATGTTGTCAATTTGAA 2044
 Db 1551 CTGTAAATAGTATATATTTTGGCCCTTACAAATACCGTAACAAATGTTGTCAATTTGAA 1610
 Qy 2045 ATACTTAATGCAAGTAACAATGCAATGCTTGGAAATTTGGAAATGATGTTATCTTTG 2104
 Db 1611 ATACTTAATGCAAGTAACAATGCAATGCTTGGAAATTTGGAAATGATGTTATCTTTG 1670
 Qy 2105 AGAAGCAAAATATGTTGCAATTAATGCTTGTATTTCTATATCAAGAAATGATGAAAG 2164
 Db 1671 AGAAGCAAAATATGTTGCAATTAATGCTTGTATTTCTATATCAAGAAATGATGAAAG 1730
 Qy 2165 TTCTCAAAACCTGTTTAAAGTACTGTTAGAGGAGCGGTTTGGAGAGACCATTTGCA 2224
 Db 1731 TTCTCAAAACCTGTTTAAAGTACTGTTAGAGGAGCGGTTTGGAGAGACCATTTGCA 1790
 Qy 2225 TCGCTGTCCAAGTGTCTTGTATTAAGTCTTTAAATGAGAGGCTAATCTCAAAATAC 2284
 Db 1791 TCGCTGTCCAAGTGTCTTGTATTAAGTCTTTAAATGAGAGGCTAATCTCAAAATAT 1850
 Qy 2285 TTTTAACTGCAATTTCTATTAATTAATGGGCAAGTATGCTCTTAC 2331
 Db 1851 TTTTAACTGCAATTTCTATTAATTAATTAATGGGCAAGTATGCTCTTAC 1897

RESULT 5
 US-09-724-676A-12752
 / Sequence 12752, Application US/09724676A
 / GENERAL INFORMATION:
 / APPLICANT: Comugen LTD
 / TITLE OF INVENTION: Variants of alternative splicing
 / FILE REFERENCE: 129181.4 Comugen
 / CURRENT APPLICATION NUMBER: US/09/724,676A
 / NUMBER OF SEQ ID NOS: 97222
 / SOFTWARE: PatentIn version 3.2
 / SEQ ID NO 12752
 / LENGTH: 1943
 / TYPE: DNA
 / ORGANISM: Homo sapiens
 / US-09-724-676A-12752
 Query Match 78.7%; Score 1833.4; DB 5; Length 1943;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1845; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 486 GTGGATGGAACAAGTGAAGATGCGACCCGATGATTAATGATCTTATATGCC 545
 Db 51 GTGGATGGAACAAGTGAAGATGCGACCCGATGATTAATGATCTTATATGCC 110
 Qy 546 ATCTCAAGCTGGGGCCAGGGGAAAAGTATGAATTTGACGACGACGACCTGCC 605
 Db 111 ATCTCAAGCTGGGGCCAGGGGAAAAGTATGAATTTGACGACGACGACCTGCC 170
 Qy 606 AGCGTGTGTCATGTATGTGAGATGATCAAGAGAAGAGAGAAAGAAATGGAA 665
 Db 171 AGCGTGTGTCATGTATGTGAGATGATCAAGAGAAGAGAGAAAGAAATGGAA 230
 Qy 666 AGAATGAGGAGACCTAAGCCAAAATATATCAGACGAGGCGGAGTAACGCCGATC 725
 Db 231 AGAATGAGGAGACCTAAGCCAAAATATATCAGACGAGGCGGAGTAACGCCGATC 290
 Qy 726 CACCTCACTGAATCTGGACGCGACGAGACGCAATTCATATCTACCGGAGAAAT 785
 Db 291 CACCTCACTGAATCTGGACGCGACGAGACGCAATTCATATCTACCGGAGAAAT 350
 Qy 786 CTTTAACTGTGAAGTGTGCTGACCACTTCCGAGGTTGGAGCGGAGATCGGGGTG 845
 Db 351 CTTTAACTGTGAAGTGTGCTGACCACTTCCGAGGTTGGAGCGGAGATCGGGGTG 410
 Qy 846 GCAGAAATCCCGATTCATGTTGCTCAGAAAGAAATCAAGGCCGTGTCCCTTGTCTAAT 905
 Db 411 GCAGAAATCCCGATTCATGTTGCTCAGAAAGAAATCAAGGCCGTGTCCCTTGTCTAAT 470
 Qy 906 GCTGACACACAGTACTGTTCATGCGCACCCGGGAAATGACTGGGCCATATCAGATTG 965
 Db 471 GCTGACACACAGTACTGTTCATGCGCACCCGGGAAATGACTGGGCCATATCAGATTG 530
 Qy 966 TGGTATGCAAGAAGACATTTGGAGCTGTCTTGAAACAGATTAATGATAGTGTG 1025
 Db 531 TGGTATGCAAGAAGACATTTGGAGCTGTCTTGAAACAGATTAATGATAGTGTG 590
 Qy 1026 TACTTGTCTTTTCTGTGAGTGTCTGTCTGTGCAAGGGCAGGTTGATCAGCTCAG 1085
 Db 591 TACTTGTCTTTTCTGTGAGTGTCTGTCTGTGCAAGGGCAGGTTGATCAGCTCAG 650
 Qy 1086 GAGAGAGCTTCTGTTCATATGAGCTGACAGGGCCACTCTCTACTGCTAGAGAGGT 1145
 Db 651 GAGAGAGCTTCTGTTCATATGAGCTGACAGGGCCACTCTCTACTGCTAGAGAGGT 710
 Qy 1146 ACAACAGGAAGCGGCTAGTGCAGAGAGGTGTGAAAACAGCAGCAATGCAATGTGAAA 1205
 Db 711 ACAACAGGAAGCGGCTAGTGCAGAGAGGTGTGAAAACAGCAGCAATGCAATGTGAAA 770
 Qy 1206 TTGTAGGTTTCTTCTTCCCTCATGTTCTCATGTTTGTGATGTATTTACTGATTTA 1265

```
Db 771' TTGTAGCGTTCTCTTCTCCCTCATGTTCTCATGTTTGTGCAATGATATTA 830
Qy 1266 CAAGACTAACCTTTGTTCTGATATATAAGTTACACCGTTGTTGTTTATACATCTTTGGGAA 1325
Db 831 CAAGACTAACCTTTGTTCTGATATATAAGTTACACCGTTGTTGTTTATACATCTTTGGGAA 890
Qy 1326 GCAGGAAAGCGTTTGGAAACGATACACCTTTCCAGATTCTCGACTCTTT 1385
Db 891 GCAGGAAAGCGTTTGGAAACGATACACCTTTCCAGATTCTCGACTCTTT 950
Qy 1386 GCAACAGCACTTGTCTGCGGAACCTCTTCTCGGAATGCATTCACACGATCCCAACCGT 1445
Db 951 GCAACAGCACTTGTCTGCGGAACCTCTTCTCGGAATGCATTCACACGATCCCAACCGT 1010
Qy 1446 GCAACGTTGATCTGCTTTTGGAAAGAGTTGATCTGAAATTCCTCTGTGAATTTA 1505
Db 1011 GCAACGTTGATCTGCTTTTGGAAAGAGTTGATCTGAAATTCCTCTGTGAATTTA 1070
Qy 1506 GCTTATACAAATTCAGAGATAGCAGTTTCACTGCGCACTTTTGTGTTGTTTCTAT 1625
Db 1071 GCTTATACAAATTCAGAGATAGCAGTTTCACTGCGCACTTTTGTGTTGTTTCTAT 1565
Qy 1626 ACATGAATCATAGCCAAACCTTTTGG- AAACCTGTTGTTGATGTTGTTCTAT 1684
Db 1191 ACATGAATCATAGCCAAACCTTTTGGAAACCTTTTGTGATGTTGTTCTAT 1250
Qy 1685 ACCCCAGAGACATCAAGATACACTTGTAAATAAGCTCATAGCATATATTCATACCTG 1744
Db 1251 ACCCCAGAGACATCAAGATACACTTGTAAATAAGCTCATAGCATATATTCATACCTG 1310
Qy 1745 TTGTACACTTGGGTGAAAGTATGGAGTGGAGCTTAAGATGATTAACCTACCTGTA 1804
Db 1311 TTGTACACTTGGGTGAAAGTATGGAGTGGAGCTTAAGATGATTAACCTACCTGTA 1370
Qy 1805 ATCATATGTTGTAGGAAAGCTGTTCCCATGCTTAACAGCACTTGAATCAAGCATGTC 1864
Db 1371 ATCATATGTTGTAGGAAAGCTGTTCCCATGCTTAACAGCACTTGAATCAAGCATGTC 1430
Qy 1865 AAGTGGATAGTATCTGTGGCGATATGAGAGGATGCGATGCTTTCCCATTCATCC 1924
Db 1431 AAGTGGATAGTATCTGTGGCGATATGAGAGGATGCGATGCTTTCCCATTCATCC 1490
Qy 1925 TGATGAAATGTTATAGTTAAATGTTGTTTCTTCTAGTTGTAATGTTGATG 1984
Db 1491 TGATGAAATGTTATAGTTAAATGTTGTTTCTTCTAGTTGTAATGTTGATG 1550
Qy 1985 CTGGTAAATAGGTTATATATTTTGGCTTTACAAATACCGTAACAATGTTGTCATTTGAA 2044
Db 1551 CTGGTAAATAGGTTATATATTTTGGCTTTACAAATACCGTAACAATGTTGTCATTTGAA 1610
Qy 2045 ATACTTAATGCCAAGTAAACATGCTTTTGGAAATTTGGAAGATGTTTATTTCTTTG 2104
Db 1611 ATACTTAATGCCAAGTAAACATGCTTTTGGAAATTTGGAAGATGTTTATTTCTTTG 1670
Qy 2105 AGAAGCAAAATATGTTTGCATTAATGCTTTGATGTTGTTGTTGTTGTTGTTGTTGTTG 2164
Db 1671 AGAAGCAAAATATGTTTGCATTAATGCTTTGATGTTGTTGTTGTTGTTGTTGTTGTTG 1730
Qy 2165 TTCTCAACCTGTTTACGGTACTTGGTAAGAGGAGCGCGTTTGGGAGAGACCATTCGA 2224
Db 1731 TTCTCAACCTGTTTACGGTACTTGGTAAGAGGAGCGCGTTTGGGAGAGACCATTCGA 1790
Qy 2225 TCGCTGTCCAAGTGTCTTGTGTTAGTGTCTTTAACTGGAGAGGCTAACCTCAAAATAC 2284
Db 1791 TCGCTGTCCAAGTGTCTTGTGTTAGTGTCTTTAACTGGAGAGGCTAACCTCAAAATAT 1850
Qy 2285 TTTTCTTAACTGATCTATATAAATGGGACACGATGCTCCTTAC 2331
Db 1851 TTTTCTTAACTGATCTATATAAATGGGACACGATGCTCCTTAC 1897
```

RESULT 6

```
US-09-724-676-12756
; Sequence 12756, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: CompuGen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 CompuGen
; CURRENT APPLICATION NUMBER: US/09/724,676
; NUMBER OF SEQ ID NOS: 2000-11-28
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12756
; LENGTH: 2050
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (7)-(7)
; OTHER INFORMATION: n is a,c,g, or t
; NAME/KEY: misc feature
; LOCATION: (47)-(47)
; OTHER INFORMATION: n is a,c,g, or t
US-09-724-676-12756
```

Query Match 67.6%; Score 1575.4; DB 5; Length 2050;
Best Local Similarity 84.2%; Pred. No. 0;
Matches 1963; Conservative 0; Mismatches 2; Indels 367; Gaps 2;

```
Qy 1 TTTTTCCTTCCCGAGGAGTGGGGCTGGCCCTTACTGCTTTTATAGCACCAGCTCAAGA 60
Db 39 TTTTTCCTTCCCGAGGAGTGGGGCTGGCCCTTACTGCTTTTATAGCACCAGCTCAAGA 98
Qy 61 AGGAACCTACAGCTCTTTGGAAAGGAATCTCACTAGGGGCTTGACTGCTGGGCTCTAG 120
Db 99 AGGAACCTACAGCTCTTTGGAAAGGAATCTCACTAGGGGCTTGACTGCTGGGCTCTAG 158
Qy 121 CGCTTTCACCTGTAGAAAGCAAGATGCAATTTAGAAAATTTTAACTACAGTTTATAGCTCCC 180
Db 159 CGCTTTCACCTGTAGAAAGCAAGATGCAATTTAGAAAATTTTAACTACAGTTTATAGCTCCC 218
Qy 181 TGATTGCTGTGTGGCAAAACAGTGTATCTTTCAGCGAAAGTGAACCCAGGCGCAAAATTTG 240
Db 219 TGATTGCTGTGTGGCAAAACAGTGTATCTTTCAGCGAAAGTGAACCCAGGCGCAAAATTTG 278
Qy 241 AGTCCCTCTTTAGGACCTATGACAAGGACATCACCTTTTCAGTATTTTAAAGAGCTTCAAC 300
Db 279 AGTCCCTCTTTAGGACCTATGACAAGGACATCACCTTTTCAGTATTTTAAAGAGCTTCAAC 338
Qy 301 GAGTCAGAAATAAATTTTCAGCAACCCCTTCTCCGAGGAGATGCCAGGCTCCAGCTGCATA 360
Db 339 GAGTCAGAAATAAATTTTCAGCAACCCCTTCTCCGAGGAGATGCCAGGCTCCAGCTGCATA 398
Qy 361 AGACTGAGTTTCTGGAAAGGAATGAAATTTATTTTGTCTCAGACCTTACACATAGGAA 420
Db 399 AGACTGAGTTTCTGGAAAGGAATGAAATTTATTTTGTCTCAGACCTTACACATAGGAA 458
Qy 421 GCTCAGACCTTGGCTCCGCAAAATTCAGACAAGCAGTTTCTGATCTCCCTCCGCGCTCTC 480
Db 459 GCTCAGACCTTGGCTCCGCAAAATTCAGACAAGCAGTTTCTGATCTCCCTCCGCGCTCTC 518
Qy 481 CGCCAGTGGGATGGAAACCAAGTGGAAAGATGCCACCCAGCTCATTAACATGATCTCTTAT 540
Db 519 CGCCAGTGGGATGGAAACCAAGTGGAAAGATGCCACCCAGCTCATTAACATGATCTCTTAT 578
Qy 541 ATGCCATCTCCAAGCTGGGGCCAGGGGAAAGTATGAATTCACGCGAGGACTGCACCA 600
Db 579 ATGCCATCTCCAAGCTGGGGCCAGGGGAAAGTATGAATTCACGCGAGGACTGCACCA 638
Qy 601 CTCCCAGGCTGGTGTCCATGATGTGAGAGTGATCAAGAGAGGAGGAAAGAGGAA 660
```


Db 39* TTTTNTCCAGGAGTGGGCTGGCCCTTACTGCTTTTATAAGCACCAGCTCAAGA 98
Qy 61 AGGAACCTACAGCCTCTTTGGAAAGAAATCTCACTAGGGGCTTGACTGCTGGGTCTGTAG 120
Db 99 AGGAACCTACAGCCTCTTTGGAAAGAAATCTCACTAGGGGCTTGACTGCTGGGTCTGTAG 158
Qy 121 CGCTTCTACTGTAGAAAGCAAGATGCATTTTAGAACTTTTAACTACAGTTTGTAGCTCCC 180
Db 159 CGCTTCTACTGTAGAAAGCAAGATGCATTTTAGAACTTTTAACTACAGTTTGTAGCTCCC 218
Qy 181 TGATTTGCTGTGTGGCAACAGTGTATCTTTCAGCAAGATGAAACAGGCGCAAAATTTG 240
Db 219 TGATTTGCTGTGTGGCAACAGTGTATCTTTCAGCAAGATGAAACAGGCGCAAAATTTG 278
Qy 241 AGTCCCTCTTTAGGACGTATGCAAGGACATCACCTTTTCAGTATTTTAAAGAGCTTCAAAC 300
Db 279 AGTCCCTCTTTAGGACGTATGCAAGGACATCACCTTTTCAGTATTTTAAAGAGCTTCAAAC 338
Qy 301 GAGTCAGATAAACTTCAGCAACCCCTTCTCCGAGCAGATGCCAGGCTCCAGTGCATA 360
Db 339 GAGTCAGATAAACTTCAGCAACCCCTTCTCCGAGCAGATGCCAGGCTCCAGTGCATA 398
Qy 361 AGACTGAGTTTCTGGGAAAGAAATGAAATGATATTTTGTCTCAGACCTTACACATAGGAA 420
Db 399 AGACTGAGTTTCTGGGAAAGAAATGAAATGATATTTTGTCTCAGACCTTACACATAGGAA 458
Qy 421 GCTCAGACCTGGCTCCGCCAAATCCAGCAAGCAGTTTCTGATCTCCCTCCGCCCTCTC 480
Db 459 GCTCAGACCTGGCTCCGCCAAATCCAGCAAGCAGTTTCTGATCTCCCTCCGCCCTCTC 518
Qy 481 CGCCAGTGGGATGAAACAAAGTGGAGATGGACCCCGAGTCATAAATGATGATCTCTTAT 540
Db 519 CGCCAGTGGGATGAAACAAAGTGGAGATGGACCCCGAGTCATAAATGATGATCTCTTAT 578
Qy 541 ATGCCATCTCOAGCTGGGSCCAGGGGAAAGTATGAATTCAGCAGCAGGACTGACACCA 600
Db 579 ATGCCATCTCOAGCTGGGSCCAGGGGAAAGTATGAATTCAGCAGCAGGACTGACACCA 638
Qy 601 CTCCAGCGTGGTCCATGATGTAGAGTGATCAAGAGAAAGGAGGAGGAA 660
Db 639 CTCCAGCGTGGTCCATGATGTAGAGTGATCAAGAGAAAGGAGGAGGAA 698
Qy 661 TGAAGAAATGAGGAGACCTTAAGCCAAAATATTCAGACCCAGGAGCGGAGTACACGC 720
Db 699 TGAAGAAATGAGGAGACCTTAAGCCAAAATATTCAGACCCAGGAGCGGAGTACACGC 758
Qy 721 CGATCCACCTCAGCTGAACTGGCAGCGGACGAGGACGCAATCCAAATCATACTCAGCGGA 780
Db 759 CGATCCACCTCAGCTGAACTGGCAGCGGACGAGGACGCAATCCAAATCATACTCAGCGGA 818
Qy 781 GGAATCTTTTACTGTGAGGTGGCTGGTCACGACTTCTTCGGAGTGGCGGAGATCG 840
Db 819 GGA----- 822
Qy 841 GGGTGGCAGAAATCCAGTTTCTATGTTGCTCAGAGAGAAATCAAGCGGTGCTCCCTTGT 900
Db 823----- 822
Qy 901 CTAACTGCTGCACACAGTTACTGTTCAATGGCACCCGGGAAATGACTTGGGCGCAATCACTGA 960
Db 823----- 822
Qy 961 GTTTGTGTGATGCGACAGGACATTTGGGACTGTCTTGAGAAACAGATAATGATGTG 1020
Db 823----- 822
Qy 1021 TTTTGTACTTGTCTTTCTGTAGGTTCTGTCTGTGCGAAGGGCAGGTGATCAGTGAG 1080
Db 823----- 822
Qy 1081 CTCAGGAGAGACTTCCGTCTTCTAAGTGGCTGAGGGGCGCACTCTCTACTGTGTAGGAA 1140
Db 823----- 822

Qy 1141 GAGGTACCAACAGGAAGCCGCTAGTGTGACAGAGGTTGTGAAACAGCAGCAATGCAATGT 1200
Db 823----- 872
Qy 1201 GGAATTTGTAGCGTTTCCCTTCTCCCTCATCTTCTCATGTTTGTGTCATGTATATTACTG 1260
Db 873 GGAATTTGTAGCGTTTCCCTTCTCCCTCATCTTCTCATGTTTGTGTCATGTATATTACTG 932
Qy 1261 ATTTCAAGACTAACCTTTGTTTCGTATATAAAGTTTACACCGTGTGTTTACATCTTTT 1320
Db 933 ATTTCAAGACTAACCTTTGTTTCGTATATAAAGTTTACACCGTGTGTTTACATCTTTT 992
Qy 1321 GGGAGCCAGGAAGCGTTTGGAAAAAGTATCACCTTTCCCAAGATTCCTCGGATTCCTGCAC 1380
Db 993 GGGAGCCAGGAAGCGTTTGGAAAAAGTATCACCTTTCCCAAGATTCCTCGGATTCCTGCAC 1052
Qy 1381 TCTTTGCAACAGCAGCTTGTCTGCGGAACTCTTCTCTGGAATGCAATTCACCTCAGCATCCCA 1440
Db 1053 TCTTTGCAACAGCAGCTTGTCTGCGGAACTCTTCTCTGGAATGCAATTCACCTCAGCATCCCA 1112
Qy 1441 ACCGTGCAACGTGTAACTTTGCTCTTTTGCAAAAGAGTTGATCTGAAATTCCTCTGTAGA 1500
Db 1113 ACCGTGCAACGTGTAACTTTGCTCTTTTGCAAAAGAGTTGATCTGAAATTCCTCTGTAGA 1172
Qy 1501 ATTTAGCTTATACAAATTCAGAGATAGCAGTTTCTCTGCCAACTTTTGTAGTGGGTGAGAAA 1560
Db 1173 ATTTAGCTTATACAAATTCAGAGATAGCAGTTTCTCTGCCAACTTTTGTAGTGGGTGAGAAA 1232
Qy 1561 TTTTAGTTTAGGTGTTGGGATCGGACCTCAGTTTCTCTGTTCTTTTATGTGCTGGT 1620
Db 1233 TTTTAGTTTAGGTGTTGGGATCGGACCTCAGTTTCTCTGTTCTTTTATGTGCTGGT 1292
Qy 1621 TCTATACATGAATTCATGCCAAAACTTTTGTGG-AAA-CTCTGTTGTGAGATAGTTGTT 1679
Db 1293 TCTATACATGAATTCATGCCAAAACTTTTGTGGAAAACTTGTGTTGAGATAGTTGTT 1352
Qy 1680 CTTTACCCCGCAGGAGCATCAAGATACACTTGTAAATTAAGCTGATAGCATATATTCAT 1739
Db 1353 CTTTACCCCGCAGGAGCATCAAGATACACTTGTAAATTAAGCTGATAGCATATATTCAT 1412
Qy 1740 ACTGTTGTACACTTGGGTGAAAAAGTATGGCAGTGGGAGACTAAGATGTATTAACCTACC 1799
Db 1413 ACTGTTGTACACTTGGGTGAAAAAGTATGGCAGTGGGAGACTAAGATGTATTAACCTACC 1472
Qy 1800 TGTGAATCATATGTTGTAGGAAAGCTGTTCCCATGTCTAACAGGACTTGAATTCAGAGC 1859
Db 1473 TGTGAATCATATGTTGTAGGAAAGCTGTTCCCATGTCTAACAGGACTTGAATTCAGAGC 1532
Qy 1860 ATGTCAAAGTGGATAGTAGTCTGTGGCGATATGAGAGGGATGCGAGTCCCTTCCCAATTC 1919
Db 1533 ATGTCAAAGTGGATAGTAGTCTGTGGCGATATGAGAGGGATGCGAGTCCCTTCCCAATTC 1592
Qy 1920 ATTCTGATGGAATGTTTACTAGGTTAACTTTGTAATTTTTTCTAGTGTGAATGTG 1979
Db 1593 ATTCTGATGGAATGTTTACTAGGTTAACTTTGTAATTTTTTCTAGTGTGAATGTG 1652
Qy 1980 TATGCTCTGTAATAGGTATATATTTGGCCCTTACATACCGTAACTAGTTTGTCAAT 2039
Db 1653 TATGCTCTGTAATAGGTATATATTTGGCCCTTACATACCGTAACTAGTTTGTCAAT 1712
Qy 2040 TTGAATTAATTAATGCAAGTAACTATGCTTTGGAAATTTGGAAGATGCTTTTAT 2099
Db 1713 TTGAATTAATTAATGCAAGTAACTATGCTTTGGAAATTTGGAAGATGCTTTTAT 1772
Qy 2100 CTTTGAAGCAAAATATGTTTGCATTAATGCTTTGATTTTCAATCAAGAAATGAT 2159
Db 1773 CTTTGAAGCAAAATATGTTTGCATTAATGCTTTGATTTTCAATCAAGAAATGAT 1832
Qy 2160 GAACGTTCTCAACCCCTGTTTACGTTACTGTTAGAGGGGAGCGGTTTGGGAGACCA 2219
Db 1833 GAACGTTCTCAACCCCTGTTTACGTTACTGTTAGAGGGGAGCGGTTTGGGAGACCA 1892

QY 2220 TTGCATCGCTGTCGAAGTCTTCTTGAAGTCTTTAACTGAGAGGCTAACTCA 2279
 DB 1893 TTGCATCGCTGTCGAAGTCTTCTTGAAGTCTTTAACTGAGAGGCTAACTCA 1952
 QY 2280 AATACCTTTTAACTGATCTATATATATATGAGGACAGATAGCTCTTAC 2331
 DB 1953 AATATTTTAACTGATCTATATATATATGAGGACAGATAGCTCTTAC 2004

RESULT 8

US-09-724-676-12748

Sequence 12748, Application US/09724676

GENERAL INFORMATION:

APPLICANT: Comugen LTD

TITLE OF INVENTION: Variants of alternative splicing

FILE REFERENCE: 129181.4 Comugen

CURRENT APPLICATION NUMBER: US/09/724,676

CURRENT FILING DATE: 2000-11-28

NUMBER OF SEQ ID NOS: 97222

SOFTWARE: PatentIn version 3.2

SEQ ID NO 12748

LENGTH: 1931

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc_feature

LOCATION: (11)..(11)

OTHER INFORMATION: n is a,c,g, or t

FEATURE:

NAME/KEY: misc_feature

LOCATION: (18)..(18)

OTHER INFORMATION: n is a,c,g, or t

FEATURE:

NAME/KEY: misc_feature

LOCATION: (23)..(23)

OTHER INFORMATION: n is a,c,g, or t

US-09-724-676-12748

Query Match 63.2%; Score 1472.8; DB 5; Length 1931;
 Best Local Similarity 83.4%; Pred. No. 0;
 Matches 1861; Conservative 0; Mismatches 3; Indels 367; Gaps 2;

QY 102 TGACTGCGTGGTCTGTAGCGCTTCACTGTAAAGAAAGCAAGATGATTTAGAACTTT 161
 DB 21 TATCTGGGTGGTGTGATAGCGCTTCACTGTAAAGAAAGCAAGATGATTTAGAACTTT 80
 QY 162 AACTACAGTTTACCTCCCTGATTTGCTGTGGCAAAACAGTGTATCTTCAGGGAAGT 221
 DB 81 AACTACAGTTTACCTCCCTGATTTGCTGTGGCAAAACAGTGTATCTTCAGGGAAGT 140
 QY 222 GAAACGAGGGGCAAAATTTAGATCCCTCTTTAGACGTATGACAAGACATCACCTTTG 281
 DB 141 GAAACGAGGGGCAAAATTTAGATCCCTCTTTAGACGTATGACAAGACATCACCTTTG 200
 QY 282 TATTTAAGAGCTTCAACGAGTCAAGATTAACCTTCAGCAACCCCTTCTCCGACGAT 341
 DB 201 TATTTAAGAGCTTCAACGAGTCAAGATTAACCTTCAGCAACCCCTTCTCCGACGAT 260
 QY 342 GCCAGGCTCCAGTGCATTAAGTGAAGTTCTGGGAAAGGAAATGAAGTATATTTTGGT 401
 DB 261 GCCAGGCTCCAGTGCATTAAGTGAAGTTCTGGGAAAGGAAATGAAGTATATTTTGGT 320
 QY 402 CAGACCTTACATAGAAAGTCAACAGCTGCTCCGCCAAATCCAGACAGAGCTTTCTG 461
 DB 321 CAGACCTTACATAGAAAGTCAACAGCTGCTCCGCCAAATCCAGACAGAGCTTTCTG 380
 QY 462 ATCTCCCTCCCTGCTCTCCGCACTGGAGTGAAGAAACAAGTGAAGATGCAACCCGATC 521
 DB 381 ATCTCCCTCCCTGCTCTCCGCACTGGAGTGAAGAAACAAGTGAAGATGCAACCCGATC 440
 QY 522 ATTAACATATATCTTTATATGCGCATCTCAAGCTGGGGCCAGGGGAAAGATGAATTTG 581
 DB 441 ATTAACATATATCTTTATATGCGCATCTCAAGCTGGGGCCAGGGGAAAGATGAATTTG 500

QY 582 CACGACCGCATGACACACCTCCACAGTGGTGTCTCATGTATGTAGATGATCAAGAG 641
 DB 501 CACGACCGCATGACACACCTCCACAGTGGTGTCTCATGTATGTAGATGATCAAGAG 560
 QY 642 AAGAGAGAAAGAGGAAATGAGAAATGAGAGACCTTAAGCAAAAATTTATCCAGAC 701
 DB 561 AAGAGAGAAAGAGGAAATGAGAAATGAGAGACCTTAAGCAAAAATTTATCCAGAC 620
 QY 702 AGAGCCCGAGTACACGCCGATCCACCTGAGTGAACCTGGCAGCGGACGAGACGAT 761
 DB 621 AGAGCCCGAGTACACGCCGATCCACCTGAGTGAACCTGGCAGCGGACGAGACGAT 680
 QY 762 CCAATATATCTACGCGGAGAACTTTTACTGTGAGGAGTGGTCTGACGACTTCTTG 821
 DB 681 CCAATATATCTACGCGGAGAA----- 703
 QY 822 GAGTGGCAGCGGAGATCGGGGTGGCAAAATCCAGTTCATGTGTCTCAGAGAGATC 881
 DB 704 ----- 703
 QY 882 AAGCCGTCCTCTGTTCTAATGCTGCAACACAGTTACTGTTCAATGCAACCGGAA 941
 DB 704 ----- 703
 QY 942 GACTTGGCCATCATGAGTTGTGTGATCGCAAGACATTTGGACGTCTTGAG 1001
 DB 704 ----- 703
 QY 1002 AAAACAGATATATATATGTTTGTGACTGTTCTTTCTGTAGTCTCTGTGCGCAA 1061
 DB 704 ----- 703
 QY 1062 GGGCAGTTGATCACTGAGCTCAGAGAGAGCTTCTGTTCTTAAGTGGCTGAGGGC 1121
 DB 704 ----- 703
 QY 1122 CACTCTACTGTAGAGAAAGATACACAGAAAGCGGCTAGTGCAGAGAGTGTGAA 1181
 DB 704 -----AGAAAGCGGCTAGTGCAGAGAGTGTGAA 734
 QY 1182 AACAGAGCAATGCAATGTGAAATTTAGCGTTCTTCTTCTCCATGTTCTCATGT 1241
 DB 735 AACAGAGCAATGCAATGTGAAATTTAGCGTTCTTCTTCTCCATGTTCTCATGT 794
 QY 1242 TTGTGATGATATATATGATTTACAGACTTACCTTTGCTGATATTAAGTTACACG 1301
 DB 795 TTGTGATGATATATATGATTTACAGACTTACCTTTGCTGATATTAAGTTACACG 854
 QY 1302 TTGTGTTTATACCTTTTGGGAACCGAGAAAGGCTTTGGAAAAGTATCACTTTCC 1361
 DB 855 TTGTGTTTATACCTTTTGGGAACCGAGAAAGGCTTTGGAAAAGTATCACTTTCC 914
 QY 1362 AGATTCGGAATTCGACTCTTTGCAACAGCACTTGTGGAACTCTTCTGGAATG 1421
 DB 915 AGATTCGGAATTCGACTCTTTGCAACAGCACTTGTGGAACTCTTCTGGAATG 974
 QY 1422 CATTCATGAGATCCCAACCGTGCAGCTTAACCTTGTCTTTGCAAAAGAGTGA 1481
 DB 975 CATTCATGAGATCCCAACCGTGCAGCTTAACCTTGTCTTTGCAAAAGAGTGA 1034
 QY 1482 TCTGAATTCCTCTGTAGAAATTTAGCTTATACAAATTCAGAGAAATGACGTTCACTGCA 1541
 DB 1035 TCTGAATTCCTCTGTAGAAATTTAGCTTATACAAATTCAGAGAAATGACGTTCACTGCA 1094
 QY 1542 ACTTTAGTGGTGAAGAAATTTAGTTAGGTGTTGGATCGGACCTGATTTCTGTTG 1601
 DB 1095 ACTTTAGTGGTGAAGAAATTTAGTTAGGTGTTGGATCGGACCTGATTTCTGTTG 1154
 QY 1602 TTTCTTTATGTGTGTTCTTATATGAAATCATAGCCAAAACCTTTTGG-AACTG 1660
 DB 1155 TTTCTTTATGTGTGTTCTTATATGAAATCATAGCCAAAACCTTTTGGAAAACCTG 1214

```
QY 1661 TTGGTTGAGATAGTTGGTCTCTTTTACCCACGAAGACATCAAGATACACTTGTAAATAAA 1720
Db 1215 TTGGTTGAGATAGTTGGTCTCTTTTACCCACGAAGACATCAAGATACACTTGTAAATAAA 1274
QY 1721 GCTGATAGCATATATTCATACACTTGTGTACACTTGGGTGAAAGATATGCGAGTGGGAGAC 1780
Db 1275 GCTGATAGCATATATTCATACACTTGTGTACACTTGGGTGAAAGATATGCGAGTGGGAGAC 1334
QY 1781 TAAGATGTATTAACCTACTGCTGGAATCATATGTTGTAGGAAAGCTGTCCCATGTCTAA 1840
Db 1335 TAAGATGTATTAACCTACTGCTGGAATCATATGTTGTAGGAAAGCTGTCCCATGTCTAA 1394
QY 1841 CAGGACTTGAATTCAAAGCATGCAAGTGGATAGTAGATCTGTGCGGATATGAGAGGAT 1900
Db 1395 CAGGACTTGAATTCAAAGCATGCAAGTGGATAGTAGATCTGTGCGGATATGAGAGGAT 1454
QY 1901 CGAGTGCCTTCCCATTCATTCCTGATGGAATTTGTATACCTAGCTTAAACATTTGTAATT 1960
Db 1455 CGAGTGCCTTCCCATTCATTCCTGATGGAATTTGTATACCTAGCTTAAACATTTGTAATT 1514
QY 1961 TTTTCTAGTTGTAATGTATGTCTGTTAAATAGGTATTATATTTTGGCCTTACAATAC 2020
Db 1515 TTTTCTAGTTGTAATGTATGTCTGTTAAATAGGTATTATATTTTGGCCTTACAATAC 1574
QY 2021 CGTAACTATGTTGTCTATTTTGAATACCTTAATGCCAAGTAAACAATGCATGCTTTGAAA 2080
Db 1575 CGTAACTATGTTGTCTATTTTGAATACCTTAATGCCAAGTAAACAATGCATGCTTTGAAA 1634
QY 2081 TTTGAGATAGTTTATCTTTTGAAGCAAAATATGTTTGCATTAATGCTTTGATTGT 2140
Db 1635 TTTGAGATAGTTTATCTTTTGAAGCAAAATATGTTTGCATTAATGCTTTGATTGT 1594
QY 2141 TCATATCAAGAAATGATGTAAGCTTCTCAACCTGTTTACGGTACTTGGTAAGAGGGA 2200
Db 1695 TCATATCAAGAAATGATGTAAGCTTCTCAACCTGTTTACGGTACTTGGTAAGAGGGA 1754
QY 2201 GCCGGTTGGGAGAGACCATTCATCTGCTCAAGTGTCTTGTAAAGTGTCTTTTAA 2260
Db 1755 GCCGGTTGGGAGAGACCATTCATCTGCTCAAGTGTCTTGTAAAGTGTCTTTTAA 1814
QY 2261 CTGGAGAGCTAACCTCAAAATACTTTTTTAACTGCACTTATAATAATGGGCACAGT 2320
Db 1815 CTGGAGAGCTAACCTCAAAATACTTTTTTAACTGCACTTATAATAATGGGCACAGT 1874
QY 2321 ATGCTCCTTAC 2331
Db 1875 ATGCTCCTTAC 1885
```

```
RESULT 9
US-09-724-676A-12748
; Sequence 12748, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12748
; LENGTH: 1931
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (11)..(11)
; OTHER INFORMATION: n is a,c,g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (18)..(18)
; OTHER INFORMATION: n is a,c,g, or t
; FEATURE:
```

```
; NAME/KEY: misc_feature
; LOCATION: (23)..(23)
; OTHER INFORMATION: n is a,c,g, or t
US-09-724-676A-12748
```

```
Query Match 63.2%; Score 1472.8; DB 5; Length 1931;
Best Local Similarity 83.4%; Pred. No. 0;
Matches 1861; Conservative 0; Mismatches 3; Indels 367; Gaps 2;
```

```
QY 102 TGACTCGTGGGTCTGTAGCGCTTTTCACTGTAGAAAGCAAGATGCATTTTAGAAACTTT 161
Db 21 TANTCTCGTGGGTCTGTAGCGCTTTTCACTGTAGAAAGCAAGATGCATTTTAGAAACTTT 80
QY 162 AACTACAGTTTTAGCTTCCCTGATTCCTGTGTGGCAAAACAGTGATATCTTCAGCGAAAGT 221
Db 81 AACTACAGTTTTAGCTTCCCTGATTCCTGTGTGGCAAAACAGTGATATCTTCAGCGAAAGT 140
QY 222 GAAACAGGCGCCAAATTTTGAAGTCCCTCTTTAGACGTATGACAAGGACATCACCTTTTCAG 281
Db 141 GAAACAGGCGCCAAATTTTGAAGTCCCTCTTTAGACGTATGACAAGGACATCACCTTTTCAG 200
QY 282 TATTTTAAGAGCTTCAAAACGAGTCAGATAAACTTCAGCAACCCCTTCTCCGCGACAGAT 341
Db 201 TATTTTAAGAGCTTCAAAACGAGTCAGATAAACTTCAGCAACCCCTTCTCCGCGACAGAT 260
QY 342 GCAGGCTCCAGCTGCATTAAGACTGAGTTTCTGGGAAAGGAAATGAAGTTATATTTTGTCT 401
Db 261 GCAGGCTCCAGCTGCATTAAGACTGAGTTTCTGGGAAAGGAAATGAAGTTATATTTTGTCT 320
QY 402 CAGACTTTACATAGGAAGCTCACACCTGCTCCGCAAAATCCAGACAAGCAGTTTCTG 461
Db 321 CAGACTTTACATAGGAAGCTCACACCTGCTCCGCAAAATCCAGACAAGCAGTTTCTG 380
QY 462 ATCTCCCTCCGCTCTCCGCGAGTGGATGGAACAAGTGGAGATGCGACCCAGTC 521
Db 381 ATCTCCCTCCGCTCTCCGCGAGTGGATGGAACAAGTGGAGATGCGACCCAGTC 440
QY 522 ATAACTATGATCTCTTATATGCCATCTCCAAGCTGGGGCCAGGGGAAAGATGAAATTG 581
Db 441 ATAACTATGATCTCTTATATGCCATCTCCAAGCTGGGGCCAGGGGAAAGATGAAATTG 500
QY 582 CACGAGCGATGACACACTCCAGCGTGGTGTCTCATGTATGTAGAGATGATCAAGAG 641
Db 501 CACGAGCGATGACACACTCCAGCGTGGTGTCTCATGTATGTAGAGATGATCAAGAG 560
QY 642 AAGGAGGAAGAGAGGAATGGAAGAAATGAGGAGACCTTAAGCCAAAATTTATCCAGACC 701
Db 561 AAGGAGGAAGAGAGGAATGGAAGAAATGAGGAGACCTTAAGCCAAAATTTATCCAGACC 620
QY 702 AGGAGCGGAGTACACCGCGATCCACCTCAGCTGGAATCGGCACGCGACGAGGACGAT 761
Db 621 AGGAGCGGAGTACACCGCGATCCACCTCAGCTGGAATCGGCACGCGACGAGGACGAT 680
QY 762 CCAATCATACTACGGGAGGAATCTTTTACTGTGGAGTGGCTGTGCTCAGACTTCTTCG 821
Db 681 CCAATCATACTACGGGAGGAATCTTTTACTGTGGAGTGGCTGTGCTCAGACTTCTTCG 703
QY 822 GAGTGGCAGCCGAGATCGGGGTGGCAGAAATCCAGTTTCATGTTGCTCAGAAGAGATC 881
Db 704 ----- 703
QY 882 AAGCCGTGTCCTTGTCTAATGTGCACACCACTTACTGTTTCATGTCACCCGGGAAT 941
Db 704 ----- 703
QY 942 GACTTGGCCCAATCACTGAGTTTGTGTGATCGCAAGGACATTTGGGACTGTCTTGAG 1001
Db 704 ----- 703
QY 1002 AAAACAGATAATGATAGTGTGTTTGTACTTCTTTCTGTTAGGTTCTGTCTGTGCCAA 1061
Db 704 ----- 703
```



```

QY 1062 GGGCAGCTTGATCAGTGAAGCTCAGAGAGAGCTTCTGTCTTAAGTGGCTCGAGGGGC 1121
Db 704 ----- 703
QY 1122 CACTCTACTGTAGAGAGATACACAGAGAGCCGCTAGTGCAGAGAGTTGTGA 1181
Db 704 ----- AGGAGCCGCTAGTGCAGAGAGTTGTGA 734
QY 1183 AACAGAGCAATGCAATGTGGAAATTTAGCGTTCTCTTCTCTCATGTCTCATGT 1241
Db 735 AACAGAGCAATGCAATGTGGAAATTTAGCGTTCTCTTCTCTCATGTCTCATGT 794
QY 1242 TTGTGATGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1301
Db 795 TTGTGATGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 854
QY 1302 TTGTGATGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1361
Db 855 TTGTGATGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 914
QY 1362 AGATTTCTGATTTCTGACTCTTTGCAACAGCATTTGCTGGGAACTCTCTGGAATG 1421
Db 915 AGATTTCTGATTTCTGACTCTTTGCAACAGCATTTGCTGGGAACTCTCTGGAATG 974
QY 1422 CATTCATCAGCATCCCAACCGTGCAACGCTGTAACCTTGCTTTTGCAAAAGAGTTGA 1481
Db 975 CATTCATCAGCATCCCAACCGTGCAACCGCTGTAACCTTGCTTTTGCAAAAGAGTTGA 1034
QY 1482 TCTGAATTCCTCTGTGAATTTAGCTTATTAATTAATTAATTAATTAATTAATTA 1541
Db 1035 TCTGAATTCCTCTGTGAATTTAGCTTATTAATTAATTAATTAATTAATTAATTA 1094
QY 1542 ACTTTTATGTTGGGAGAAATTTTATAGTGTGGATGGGAGCTGAGTCTGTTG 1601
Db 1095 ACTTTTATGTTGGGAGAAATTTTATAGTGTGGATGGGAGCTGAGTCTGTTG 1154
QY 1602 TTTCTTTATGTTGGGAGTCTTATTAATTAATTAATTAATTAATTAATTAATTA 1660
Db 1155 TTTCTTTATGTTGGGAGTCTTATTAATTAATTAATTAATTAATTAATTAATTA 1214
QY 1661 TTGTTGATGATGTTGTTCTTTTATCCCAACGAGAGATCAATGATCACTTTAAATTA 1720
Db 1215 TTGTTGATGATGTTGTTCTTTTATCCCAACGAGAGATCAATGATCACTTTAAATTA 1274
QY 1721 GCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1780
Db 1275 GCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1334
QY 1781 TAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1840
Db 1335 TAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1394
QY 1841 CAGACCTTGAATTAACAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1900
Db 1395 CAGACCTTGAATTAACAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1454
QY 1901 GCAGGCGCTTCCCATTCATTCCTGATGATGATGATGATGATGATGATGATGATGAT 1960
Db 1455 GCAGGCGCTTCCCATTCATTCCTGATGATGATGATGATGATGATGATGATGATGAT 1514
QY 1961 TTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2020
Db 1515 TTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1574
QY 2021 CGTAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2080
Db 1575 CGTAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1634
QY 2081 TTTGGAAGTGGTTTATCTTTGAGAGCAATGTTGATTAATGCTTTGATTTG 2140
Db 1635 TTTGGAAGTGGTTTATCTTTGAGAGCAATGTTGATTAATGCTTTGATTTG 1694
QY 2141 TCATATCAAGAAATGATGAAAGTTCTCAAAACCTGTTTACGGTACTTGGTAAGAGGA 2200

```

```

Db 1695 TCATATCAAGAAATGATGAAAGTTCTCAAAACCTGTTTACGGTACTTGGTAAGAGGA 1754
QY 2201 GCGGTTTGGGAGAGACCAATTCATGCTGCTCCAAAGTCTTTCTTTTAAAGTCTTTTAA 2260
Db 1755 GCGGTTTGGGAGAGACCAATTCATGCTGCTCCAAAGTCTTTCTTTTAAAGTCTTTTAA 1814
QY 2261 CTGAGAGGCTAACTTCAAAATCTTTTAACTGATCTTATTAATTAATGAGGCACT 2320
Db 1815 CTGAGAGGCTAACTTCAAAATCTTTTAACTGATCTTATTAATTAATGAGGCACT 1874
QY 2321 ATGCTCCTTAC 2331
Db 1875 ATGCTCCTTAC 1885

RESULT 10
US-09-724-676-12741
Sequence 12741, Application US/09724676
GENERAL INFORMATION:
APPLICANT: Compugen LTD
TITLE OF INVENTION: Variants of alternative splicing
FILE REFERENCE: 129181.4 Compugen
CURRENT APPLICATION NUMBER: US/09/724,676
CURRENT FILING DATE: 2000-11-28
NUMBER OF SEQ ID NOS: 97222
SOFTWARE: PatentIn version 3.2
SEQ ID NO 12741
LENGTH: 1893
TYPE: DNA
ORGANISM: Homo sapiens
US-09-724-676-12741

Query Match      58.1%; Score 1354.4; DB 5; Length 1893;
Best Local Similarity 80.8%; Pred. No. 0;
Matches 1775; Conservative 0; Mismatches 56; Indels 367; Gaps 2;

```


QY	675	AGACCTTAAGCCAAAAATTATCCAGACCCAGGAGCGGAGTACACGCCGATCCACCTCAGC	733
Db	556	AGACCTTAAGCCAAAAATTATCCAGACCCAGGAGCGGAGTACACGCCGATCCACCTCAGC	615
QY	735	TGAACCTGGCAGCGAGCGAGCGCATTCCTCAATCATATCTCACCGGAGGAATCTTTTACTG	794
Db	616	TGAACCTGGCAGCGAGCGAGCGCATTCCTCAATCATATCTCACCGGAGGAATCTTTTACTG	665
QY	795	TGAGGTGGCTGGTTCACGACTCTTCGGAGGTGGCAGCGAGATCGGGTGGCAAAATC	854
Db	666	-----	665
QY	855	CCAGTTTCATGTTGCTTCAGAGAGAAATCAAGCCCGTGCCTTTGTTCTTAATGTCGCACAC	914
Db	666	-----	665
QY	915	CAGTTACTGTTTCATGCGCACC CGGGAATGCACTTGGGCCAATCACTGAGTTTGTGTGATCG	974
Db	666	-----	665
QY	975	CACAAGGACATTTGGGACTGCTTTGAGAAAAACAGATAATGATAGTGTTTTGTTACTGTTTC	1034
Db	666	-----	665
QY	1035	TTTTCTGGTAGTTCCTGCTCTGCTGCAAGGGCAGGTTGATCAGTGAAGCTCAGGAGAGACT	1094
Db	666	-----	665
QY	1095	TGCTGTTTCTAAGTGGCTGCGAGGGCCACTCTCTACTGTTAGGAAGAGTTACCAAGGA	1154
Db	666	-----AGGA	669
QY	1155	AGCGCCCTAGTGCAGAGAGGTTGTGAAAAACAGCAGCAATGCAATGTGGAATTCGTAGCGT	1214
Db	670	AGCGCCCTAGTGCAGAGAGGTTGTGAAAAACAGCAGCAATGCAATGTGGAATTCGTAGCGT	729
QY	1215	TTCCCTTTCTCCCTCATGTTCTCATGTTTGTGATGTATPATTACTGATTTACAAGACTAA	1274
Db	730	TTCCCTTTCTCCCTCATGTTCTCATGTTTGTGATGTATPATTACTGATTTACAAGACTAA	789
QY	1275	CCTTTGTTCGTATATAAGTTACACCGTTGTGTTTACATCTTTTGGGAAGCCAGAAA	1334
Db	790	CCTTTGTTCGTATATAAGTTACACCGTTGTGTTTACATCTTTTGGGAAGCCAGAAA	849
QY	1335	GGTTTGGAAAAACGTATCACTTTCCAGATTTCTCGGATTTCTCGACTCTTTGCAACAGCA	1394
Db	850	GGTTTGGAAAAACGTATCACTTTCCAGATTTCTCGGATTTCTCGACTCTTTGCAACAGCA	909
QY	1395	CTTGCTTCCGGAACCTCTCTCGGAATGCAATCACTCAGATCCCAACCGTCAACGTTGT	1454
Db	910	CTTGCTTCCGGAACCTCTCTCGGAATGCAATCACTCAGATCCCAACCGTCAACGTTGT	969
QY	1455	AACTTGTGCTTTTGCAAAAGAGTTGATCTGAAATTCCTCTGTAGATTTAGCTTATACA	1514
Db	970	AACTTGTGCTTTTGCAAAAGAGTTGATCTGAAATTCCTCTGTAGATTTAGCTTATACA	1029
QY	1515	ATTCAGAGATAGAGTTTCACTGCGCAACTTTTGTGGGTGAGAAAATTTTGTAGTTAGGTTG	1574
Db	1030	ATTCAGAGATAGAGTTTCACTGCGCAACTTTTGTGGGTGAGAAAATTTTGTAGTTAGGTTG	1089
QY	1575	TTTGGGATCGGACCTCAGTTTCTGTGTGTTCTTTTATGTGGTGGTTCATATCATGAATC	1634
Db	1090	TTTGGGATCGGACCTCAGTTTCTGTGTGTTCTTTTATGTGGTGGTTCATATCATGAATC	1149
QY	1635	ATAGCCAAAAAATTTTTTGG-AAACTGTTGGTTGAGATAGTTGGTCTTTTATCCCAAGA	1693
Db	1150	ATAGCCAAAAAATTTTTTGGAAAACTGTTGGTTGAGATAGTTGGTCTTTTATCCCAAGA	1209
QY	1694	AGACATCAAGATACCTTGTAAATAAGCTGATAGCATATATTATACACCTGTTGTACACT	1753
Db	1210	AGACATCAAGATACCTTGTAAATAAGCTGATAGCATATATTATACACCTGTTGTACACT	1269
QY	1754	TGGGTGAAAAGTATGCGACTGCGGAGACTTAAGATGATTAACCTACCTGTGGAATCATATGT	1813

Db	1270	TGGGTGAAAGATGTCGAGTGGGAGACTAAGATGTATTAACTACCTCGTGAATCATATGT	1329
QY	1814	TGTAGGAAAAGCTGTGTCCCATCTCTAACAGGACTTGAAATCCAAGCATGTCAAGTGGATA	1873
Db	1330	TGTAGGAAAAGCTGTGTCCCATCTCTAACAGGACTTGAAATCCAAGCATGTCAAGTGGATA	1389
QY	1874	GTAGATCTGTGGCGATATGAGAGGGATGCAGTGCCTTTCCCATTCAATCCTGTATGGAAT	1933
Db	1390	GTAGATCTGTGGCGATATGAGAGGGATGCAGTGCCTTTCCCATTCAATCCTGTATGGAAT	1449
QY	1934	TGTATACTAGTGTAAACATTGTAAATTTTTTTCTAGTGTGAATGTATGTCTGGTAAAT	1993
Db	1450	TGTATACTAGTGTAAACATTGTAAATTTTTTTCTAGTGTGAATGTATGTCTGGTAAAT	1509
QY	1994	AGGTATATATTTTGGCCTTTACAATACCCTAACAAATGTTTGTTCATTTTGAATACTTAAT	2053
Db	1510	AGGTATATATTTTGGCCTTTACAATACCCTAACAAATGTTTGTTCATTTTGAATACTTAAT	1569
QY	2054	GCCAAGTAACAAATGCATGCTTTGGAAATTTGGAAGATGGTTTTATCTTTGAGAAGCAA	2113
Db	1570	GCCAAGTAACAAATGCATGCTTTGGAAATTTGGAAGATGGTTTTATCTTTGAGAAGCAA	1629
QY	2114	TATGTTTGCATTAATGCTTTGATCTTTCATATCAAGAAATTTGATGAAAGCTTCTCAAC	2173
Db	1630	TATGTTTGCATTAATGCTTTGATCTTTCATATCAAGAAATTTGATGAAAGCTTCTCAAC	1689
QY	2174	CCTGTTTACGGTACTTGTTAAGGAGGCCGTTTGGGAGAGACCATTTGCATCGCTGTCC	2233
Db	1690	CCTGTTTACGGTACTTGTTAAGGAGGCCGTTTGGGAGAGACCATTTGCATCGCTGTCC	1749
QY	2234	AAGTGTTCCTTGTGTTAAGTGCTTTTAACTTGGAGAGGCTTAACTTCAAAATACTTTTTTAA	2293
Db	1750	AAGTGTTCCTTGTGTTAAGTGCTTTTAACTTGGAGAGGCTTAACTTCAAAATACTTTTTTAA	1809
QY	2294	CTGCATTCATATAAATGGCACAGTATGTCCTTAC	2331
Db	1810	CTGCATTCATATAAATGGCACAGTATGTCCTTAC	1847

```

RESULT 11
US-09-724-676A-12741
; Sequence 12741, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: CompuGen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 CompuGen
; CURRENT APPLICATION NUMBER: US/09/724.676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12741
; LENGTH: 1893
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676A-12741

```

Query Match 58.1%; Score 1354.4; DB 5; Length 1993;
Best Local Similarity 80.8%; Pred. No. 0;
Matches 1775; Conservative 0; Mismatches 56; Indels 367; Gaps 2;

	Matches 1775;	Conservative	0;	Mismatches	56;	Indels	367;	Gaps		
Qy	135	GAAGCAGAGTGCATTTT	TAGAACTT	TAAC	TA	CAGTTTTAGCTTC	CCTGAT	TGCTGTG	134	
Db	16	GACTGCAGATGGAGG	AGGTGG	ACCT	CGAGG	AGCTGCC	AGCGCC	ACCATCGCT	75	
Qy	195	GCAAA	CAGTGATATCTT	TCACG	AAAGTG	AAAC	CAGGGCC	AAATTT	CAGTCCCTCTTTAGG	254
Db	76	CTGG	ACCGCGCTGTTCTG	TGG	AGCGCTG	CTGCGG	CCAAATTT	TTCAGTCCCTCTTTAGG	135	
Qy	255	ACGTATG	ACAAGGACAT	CACTTT	CTCAGTAT	TTTTTAAG	AGCTTCA	AAACGAGTC	GAGAAATAAC	314
Db	136	ACGTATG	ACAAGGACAT	CACTTT	CTCAGTAT	TTTTTAAG	AGCTTCA	AAACGAGTC	GAGAAATAAC	195

QY 315 TTCAGCAACCCCTTCTCCGACAGATGCGAGCTCCAGCTGCATTAAGTGTCTG 374
 Db 196 TTCAGCAACCCCTTCTCCGACAGATGCGAGCTCCAGCTGCATTAAGTGTCTG 255
 QY 375 GGAAGGAATAGTATATATTTTGTCTCAGACCTTACATAGAACTCACAAGTCTG 434
 Db 256 GGAAGGAATAGTATATATTTTGTCTCAGACCTTACATAGAACTCACAAGTCTG 315
 QY 435 CCGCCAAATCCAGACAGAGTTTCTGATCTCCCTCCGCTCCGCGAGTGGATGG 494
 Db 316 CCGCCAAATCCAGACAGAGTTTCTGATCTCCCTCCGCTCCGCGAGTGGATGG 375
 QY 495 AAACAGTGAAGATGCGAGCCCGAGTCATTAATATATATATATATATATATAT 554
 Db 376 AAACAGTGAAGATGCGAGCCCGAGTCATTAATATATATATATATATATATAT 435
 QY 555 CTGGGCGCAGGGAAGATATGATTCAGACGCGAGCTGACACCACTCCGAGCTGGT 614
 Db 436 CTGGGCGCAGGGAAGATATGATTCAGACGCGAGCTGACACCACTCCGAGCTGGT 495
 QY 615 GTCCATGATGTAGAGTATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 674
 Db 496 GTCCATGATGTAGAGTATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 555
 QY 675 AGACCTAAGCCAAATATATCCAGACGAGAGCCGAGTACACGCCGATCCACTCAGC 734
 Db 556 AGACCTAAGCCAAATATATCCAGACGAGAGCCGAGTACACGCCGATCCACTCAGC 615
 QY 735 TGAACCTGACGCGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 794
 Db 616 TGAACCTGACGCGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 665
 QY 795 TGAAGTGGCTGTGACAGACTTCTTGGAGGTGGCAGAGTGGGAGTGGAGAAATC 854
 Db 666 ----- 665
 QY 855 CCAATCATGTGCTCAGAGAGAGATCAAGCCGTGTCCTTGTCTAATGCTGACAC 914
 Db 666 ----- 665
 QY 915 CAGTACTGTTCATGACACCGGGAATGACTTGGGCAATCAGTATTTGTGTATCG 974
 Db 666 ----- 665
 QY 975 CACAAGCATTGGAGCTGTCTTGAAGAAACAGATATGATGTTTGTACTTGTTC 1034
 Db 666 ----- 665
 QY 1035 TTTCTGTAGTCTGTCTGTGCGCAAGGAGGTGATCAGTACGCTCAGAGAGAGCT 1094
 Db 666 ----- 665
 QY 1095 TCTGTCTTAAGTGGCTGCGAGGGGCACTCTTACTGTAGAGAAAGTACCACAGGA 1154
 Db 666 ----- 665
 QY 1155 AGCGCCCTAGTACAGAGAGTGTGAAAAACAGACCAATGCAATGTGAAAAATGAGCT 1214
 Db 670 AGCGCCCTAGTACAGAGAGTGTGAAAAACAGACCAATGCAATGTGAAAAATGAGCT 729
 QY 1215 TTTCTTTCTCCCTCATGTTCTCATGTTTGTGATGTAATTTACTGATTTACAAAGCTAA 1274
 Db 730 TTTCTTTCTCCCTCATGTTCTCATGTTTGTGATGTAATTTACTGATTTACAAAGCTAA 789
 QY 1275 CTTTGTCTGTATTAAGTTACACGTTGTTTTCATCTTTTGGAAAGCCAGAAA 1334
 Db 790 CTTTGTCTGTATTAAGTTACACGTTGTTTTCATCTTTTGGAAAGCCAGAAA 849
 QY 1335 GCGTTTGAAGAGTATCACTTCCAGATTCCTGATTCGACTTTTGGACAGCA 1394
 Db 850 GCGTTTGAAGAGTATCACTTCCAGATTCCTGATTCGACTTTTGGACAGCA 909
 QY 1395 CTTCGTTGGGAACTCTCTCTGAGATGCACTTCACTCAGATCCCAACCGTGGCAAGTGT 1454

Db 910 CTTCCTTGGGAACTCTCTCTGAGATGCACTCAGATCCCAACCGTGGCAAGTGT 969
 QY 1455 AACTGTGCTTTTGAAGAAAGATGATCTGAATTCCTGTGAGATTTTACCTATACA 1514
 Db 970 AACTGTGCTTTTGAAGAAAGATGATCTGAATTCCTGTGAGATTTTACCTATACA 1029
 QY 1515 ATTGAGAAATGAGTGTTCATCTGCCAATTTTATGAGGTGAGAAATTTTATGTTAGGTG 1574
 Db 1030 ATTGAGAAATGAGTGTTCATCTGCCAATTTTATGAGGTGAGAAATTTTATGTTAGGTG 1089
 QY 1575 TTTGGGATCGGACCTGCTGTTGTTTCTTTATGAGGTGTTTATACATGAATC 1634
 Db 1090 TTTGGGATCGGACCTGCTGTTGTTTCTTTATGAGGTGTTTATACATGAATC 1149
 QY 1635 ATAGCCAAAACTTTTGGG-AAAAGTGTGAGTGAATGTTTCTTTTACCCACGA 1693
 Db 1150 ATAGCCAAAACTTTTGGGAAACGTGTGTGAGTGAATGTTTCTTTTACCCACGA 1209
 QY 1694 AGACATCAAGTATACATTTGTAATTAAGCTGATAGCTATATATATATATATATATAT 1753
 Db 1210 AGACATCAAGTATACATTTGTAATTAAGCTGATAGCTATATATATATATATATATAT 1269
 QY 1754 TGGGGAAGAGTATGAGAGTGGAGAGCTAAGATGTATTAACCTGATGATCATATGT 1813
 Db 1270 TGGGGAAGAGTATGAGAGTGGAGAGCTAAGATGTATTAACCTGATGATCATATGT 1329
 QY 1814 TGTAGAAAGCTGTTCATGTCTTAACAGAGCTTGAATTCAAAGCATGTCAAGTATGATA 1873
 Db 1330 TGTAGAAAGCTGTTCATGTCTTAACAGAGCTTGAATTCAAAGCATGTCAAGTATGATA 1389
 QY 1874 GTAGATCTGTGCGCATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1933
 Db 1390 GTAGATCTGTGCGCATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1449
 QY 1934 TGTATCTAGGTAACTTTGTAATTTTCTAGTGTATGATGTCTGTAAAT 1993
 Db 1450 TGTATCTAGGTAACTTTGTAATTTTCTAGTGTATGATGTCTGTAAAT 1509
 QY 1994 AGGATATATATTTGCGCTTACATACCGTAAACATGTTTGTCAATTTTGAATTTCTAAT 2053
 Db 1510 AGGATATATATTTGCGCTTACATACCGTAAACATGTTTGTCAATTTTGAATTTCTAAT 1569
 QY 2054 GCCAAGTAACTATGATGATGTTTGAAGATGTTTATTTCTTGAAGAGCAAA 2113
 Db 1570 GCCAAGTAACTATGATGATGTTTGAAGATGTTTATTTCTTGAAGAGCAAA 1629
 QY 2114 TATGTTGATTAATATGCTTTGATTTTCAATCAAGAAATTTGATTGAAGCTTCAAAAC 2173
 Db 1630 TATGTTGATTAATATGCTTTGATTTTCAATCAAGAAATTTGATTGAAGCTTCAAAAC 1689
 QY 2174 CCTGTTTACGTAATCTTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2233
 Db 1690 CCTGTTTACGTAATCTTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1749
 QY 2234 AAGTGTCTGTGTAAGTCTTTTAACTGAGAGGCTAATCTCAAAATCTTTTAA 2293
 Db 1750 AAGTGTCTGTGTAAGTCTTTTAACTGAGAGGCTAATCTCAAAATCTTTTAA 1809
 QY 2294 CTGCATTTAATTAATATGAGGACAGTATGCTCTTAC 2331
 Db 1810 CTGCATTTAATTAATATGAGGACAGTATGCTCTTAC 1847

RESULT 12
 US-09-724-676-12745
 ; Sequence 12745, Application US/09724676
 ; GENERAL INFORMATION:
 ; APPLICANT: Compugen LTD
 ; TITLE OF INVENTION: Variants of alternative splicing
 ; FILE REFERENCE: 129181.4 Compugen
 ; CURRENT APPLICATION NUMBER: US/09/724,676
 ; CURRENT FILING DATE: 2000-11-28

; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12745
; LENGTH: 1876
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676-12745

Query Match 57.9%; Score 1349.4; DB 5; Length 1876;
Best Local Similarity 81.8%; Pred. No. 0;
Matches 1749; Conservative 0; Mismatches 21; Indels 367; Gaps 2;
QY 196 CAAACAGTGCATCTTCACGCGAAGTGAACCCAGGGCCAAATTTGAGTCCCTCTTTAGGA 255
DB 60 CAATCTGCNCTATTGGAAGAAACCAATGTTTGGGCCAAATTTGAGTCCCTCTTTAGGA 119
QY 256 CGTATGACAGGACATCACTTTTCAGTATTTTAAAGAGCTTCAACCGAGTCAAGATAACT 315
DB 120 CGTATGACAGGACATCACTTTTCAGTATTTTAAAGAGCTTCAACCGAGTCAAGATAACT 179
QY 316 TCAGCAACCCCTTCCGAGGAGATGCCAGGCTCCAGCTGCATAAGACTGAGTTTCTGG 375
DB 180 TCAGCAACCCCTTCCGAGGAGATGCCAGGCTCCAGCTGCATAAGACTGAGTTTCTGG 239
QY 376 GAAAGCAATGAAGTTATATTTGCTCAGACTTACACATAGGAAGCTCAACCTGGCTC 435
DB 240 GAAAGCAATGAAGTTATATTTGCTCAGACTTACACATAGGAAGCTCAACCTGGCTC 299
QY 436 CGCCAAATCCAGACAAGCAGTTTCTGATCTCCCTCCGCTCTCCGCTCTCCGCTGGATGGA 495
DB 300 CGCCAAATCCAGACAAGCAGTTTCTGATCTCCCTCCGCTCTCCGCTCTCCGCTGGATGGA 359
QY 496 AACAGTGGAGATGCCAGCCAGCTCAATACTATGATCTTATATGCCATCTCCAAGC 555
DB 360 AACAGTGGAGATGCCAGCCAGCTCAATACTATGATCTTATATGCCATCTCCAAGC 419
QY 556 TGGGCCAGGGGAAAGTATGAAATGACGAGGAGCTGACACCCACTCCAGCTGGTGG 615
DB 420 TGGGCCAGGGGAAAGTATGAAATGACGAGGAGCTGACACCCACTCCAGCTGGTGG 479
QY 616 TCCATGTATGTGAGGTGATCAAGAGAGGAGGAGAAATGGAAGAAATGAGGA 675
DB 480 TCCATGTATGTGAGGTGATCAAGAGAGGAGGAGAAATGGAAGAAATGAGGA 539
QY 676 GACCTAAGCCAAATTTATCCAGACGAGGAGGCGGAGTACACGCCGATCCACTCAGCT 735
DB 540 GACCTAAGCCAAATTTATCCAGACGAGGAGGCGGAGTACACGCCGATCCACTCAGCT 599
QY 736 GAATGCGACGCGACGAGGAGCGATTCCAAATCATCTACGCGGAGGAATCTTTTACTGT 795
DB 600 GAATGCGACGCGACGAGGAGCGATTCCAAATCATCTACGCGGAGGAATCTTTTACTGT 648
QY 796 GGAGTGGCTGGTCAACGACTTCTTCGAGGTGGACGCCGAGATCGGGGTGGCAAAATCC 855
DB 649 ----- 648
QY 856 CAGTTATGTTGCTCAGAGAGAAATCAAGCGCGTGGTCCCTTGTCTAATGTGTCACACC 915
DB 649 ----- 648
QY 916 AGTTACTGTTTCATGGCACCGGGAATGACTTGGGCCAATCACTGAGTTTGTGTGATCGC 975
DB 649 ----- 648
QY 976 ACAAGGACATTTGGGACTGTCTTGAGAAACACAGATAATGATAGTGTGTTGTTACTTGTCT 1035
DB 649 ----- 648
QY 1036 TTTCTGTAGGTTCTGTCTGTGCCAAGGCGAGTTGATCAGTGAGCTCAGGAGAGGCTT 1095
DB 649 ----- 648
QY 1096 CTTGTTTCTAAGTGGCTCGAGGGGCCACTCTCTACTGTAGGAAGAGGTACCCAGGA 1155

DB 649 ----- 648
QY 1156 GCAGCTAGTGCAGAGAGGTGTGAAAACAGCAGCAATGCAATGTGGAATTTGTAGCGTT 1215
DB 654 GCAGCTAGTGCAGAGAGGTGTGAAAACAGCAGCAATGCAATGTGGAATTTGTAGCGTT 713
QY 1216 TCTTTTCTCCCTCATGTTTCTCATGTTTGTGATGTATATTACTGATTTTACAAGACTAAC 1275
DB 714 TCTTTTCTCCCTCATGTTTCTCATGTTTGTGATGTATATTACTGATTTTACAAGACTAAC 773
QY 1276 CTTTGTTCGTATATAAGTATACCGTGTGTTTGTATATCTTTTGGGAAGCCAGGAAG 1335
DB 774 CTTTGTTCGTATATAAGTATACCGTGTGTTTGTATATCTTTTGGGAAGCCAGGAAG 833
QY 1336 CGTTTGGAAAACGTATCACCTTTCCAGATTCCTCGATTCCTCGATTCCTGCACTCTTTGCAAGCAGAC 1395
DB 834 CGTTTGGAAAACGTATCACCTTTCCAGATTCCTCGATTCCTCGATTCCTGCAAGCAGAC 893
QY 1396 TTGCTTTGGGAACTCTTCTCGAATGCAATTCCTCAGCATCCCAACCGTGCAACGTGTA 1455
DB 894 TTGCTTTGGGAACTCTTCTCGAATGCAATTCCTCAGCATCCCAACCGTGCAACGTGTA 953
QY 1456 ACTTGTGCTTTTGGCAAAAGAGTTCGATCTGAAATTCCTCTGAGAAATTTAGCTTATACAA 1515
DB 954 ACTTGTGCTTTTGGCAAAAGAGTTCGATCTGAAATTCCTCTGAGAAATTTAGCTTATACAA 1013
QY 1516 TTCAGAGATAGCAGTTTCTGCACTGCAACTTTTGTGGGTGAGAAATTTTGTAGTGT 1575
DB 1014 TTCAGAGATAGCAGTTTCTGCACTGCAACTTTTGTGGGTGAGAAATTTTGTAGTGT 1073
QY 1576 TTGGGATCGGACCTCAGTTTCTGTTGTTTCTTTATGTTGGTGGTTCATACATGAATCA 1635
DB 1074 TTGGGATCGGACCTCAGTTTCTGTTGTTTCTTTATGTTGGTGGTTCATACATGAATCA 1133
QY 1636 TAGCCAAAACCTTTTGGG-AAAACCTTGTGGTGTGAGATAGTTGGTTCCTTTACCCCGCA 1694
DB 1134 TAGCCAAAACCTTTTGGGAAAACCTTGTGGTGTGAGATAGTTGGTTCCTTTACCCCGCA 1193
QY 1695 GACATCAAGATACACTTGTAAATAAAGCTGATAGCATATATTCATACCTGTTGTACACTT 1754
DB 1194 GACATCAAGATACACTTGTAAATAAAGCTGATAGCATATATTCATACCTGTTGTACACTT 1253
QY 1755 GGGTGAAGATGATGCGAGTGGGACCTAAGATGATTTAACTACCTACCTGTAATCATGTT 1814
DB 1254 GGGTGAAGATGATGCGAGTGGGACCTAAGATGATTTAACTACCTACCTGTAATCATGTT 1313
QY 1815 GTAGAAAAGCTGTTCCCATGTCTTAACAGGACTTGAATTCAAAGCATGTCAAGTGGATAG 1874
DB 1314 GTAGAAAAGCTGTTCCCATGTCTTAACAGGACTTGAATTCAAAGCATGTCAAGTGGATAG 1373
QY 1875 TAGATCTGTGGCGATATGAGAGGATGCGAGTCCCTTTCCCATTCATTCCTGATGGAAT 1934
DB 1374 TAGATCTGTGGCGATATGAGAGGATGCGAGTCCCTTTCCCATTCATTCCTGATGGAAT 1433
QY 1935 GTTACTAGTGTAAACATTTGTAATTTTCTAGTTGTAAATGTTGTAATGTTGTAATGTTGTAAT 1994
DB 1434 GTTACTAGTGTAAACATTTGTAATTTTCTAGTTGTAAATGTTGTAATGTTGTAATGTTGTAAT 1493
QY 1995 GGTATTTATTTTGGCCCTTACAATAACCGTAAACATGTTTGTCAATTTTGAATACTTAATG 2054
DB 1494 GGTATTTATTTTGGCCCTTACAATAACCGTAAACATGTTTGTCAATTTTGAATACTTAATG 1553
QY 2055 CCAAGTAAACATTCAGCTTTTGGAAATTTGGAAGATGTTTATTTCTTTGAGAGCAAAAT 2114
DB 1554 CCAAGTAAACATTCAGCTTTTGGAAATTTGGAAGATGTTTATTTCTTTGAGAGCAAAAT 1613
QY 2115 ATGTTGCAATTAATGCTTTGATTTGTTTATCAAGAAATTTGATTCGAACGTTCTCAACC 2174
DB 1614 ATGTTGCAATTAATGCTTTGATTTGTTTATCAAGAAATTTGATTCGAACGTTCTCAACC 1673
QY 2175 CTGTTTACGCTACTTGGTAAAGGAGCGGTTTGGGAGAGCAATTCGATCGCTGTCCA 2234

Db 1374 TAGATCTCTGGCGATATGAGAGGATCGAGTGCCTTTCCCAATTCCTGATGAAT 1433
Qy 1935 GTTATACCTAGGTAACTTTGTAATTTTTCTAGTTGTAATGTGTATGCTCGTAAATA 1994
Db 1434 GTTATACCTAGGTAACTTTGTAATTTTTCTAGTTGTAATGTGTATGCTCGTAAATA 1493
Qy 1995 GGTATATATTTGGCCTTACAACTACCGTAACAAATGTTGTCATTTTGAATACTTTAATG 2054
Db 1494 GGTATATATTTGGCCTTACAACTACCGTAACAAATGTTGTCATTTTGAATACTTTAATG 1553
Qy 2055 CCAAGTAACAAATGATGCTTTGGAATTTTGGAAAGATGTTTATTTCTTTGAGAGCAAT 2114
Db 1554 CCAAGTAACAAATGATGCTTTGGAATTTTGGAAAGATGTTTATTTCTTTGAGAGCAAT 1613
Qy 2115 ATGTTTGCATTAAATGCTTTGATTTGTTTATTCATATCAAGAAATTCATTGAACCTTCAAAACC 2174
Db 1614 ATGTTTGCATTAAATGCTTTGATTTGTTTATTCATATCAAGAAATTCATTGAACCTTCAAAACC 1673
Qy 2175 CTGTTTACCGTACTTGGTAAGAGGAGCGGTTTGGGAGAGCAATTCGATCGCTGTCCA 2234
Db 1674 CTGTTTACCGTACTTGGTAAGAGGAGCGGTTTGGGAGAGCAATTCGATCGCTGTCCA 1733
Qy 2235 AGTGTCTTCTGTTAGTGTCTTTAACTGGAGAGGCTAACCTCAAAATACCTTTTAAAC 2294
Db 1734 AGTGTCTTCTGTTAGTGTCTTTAACTGGAGAGGCTAACCTCAAAATACCTTTTAAAC 1793
Qy 2295 TGCATTCTATAATAATGGCACAGTATGCTCTTAC 2331
Db 1794 TGCATTCTATAATAATGGCACAGTATGCTCTTAC 1830

RESULT 14

US-09-724-676-12743

; Sequence 12743, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12743
; LENGTH: 1839
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676-12743

Query Match 57.8%; Score 1348.4; DB 5; Length 1839;
Best Local Similarity 82.5%; Pred. No. 0;
Matches 1736; Conservative 0; Mismatches 1; Indels 367; Gaps 2;
Qy 229 GGGCCAAATTTGAGTCCCTCTTTAGGACGTATGACAAAGGACATCACCCTTTCAGTATTTTA 288
Db 56 GGGCCAAATTTGAGTCCCTCTTTAGGACGTATGACAAAGGACATCACCCTTTCAGTATTTTA 115
Qy 289 AGAGCTTCAACCGAGTCAGATAAATCTTCAGCAACCCCTTCTCCGAGCAGATGCCAGC 348
Db 116 AGAGCTTCAACCGAGTCAGATAAATCTTCAGCAACCCCTTCTCCGAGCAGATGCCAGC 175
Qy 349 TCCAGTGTCAATAGACTGAGTTCTTGGAAAGGAATGAAGTATATTTTGTCTCAGACT 408
Db 176 TCCAGTGTCAATAGACTGAGTTCTTGGAAAGGAATGAAGTATATTTTGTCTCAGACT 235
Qy 409 TACACATAGGAAGCTCAGACTGGCTCCGCAAAATCCAGACAGCAGTTTCTGATCTCC 468
Db 236 TACACATAGGAAGCTCAGACTGGCTCCGCAAAATCCAGACAGCAGTTTCTGATCTCC 295
Qy 469 CTCCCGCTCTCCGCGAGTGGATGGAACAGTGAAGATGCCAGCCCACTAAACT 528
Db 296 CTCCCGCTCTCCGCGAGTGGATGGAACAGTGAAGATGCCAGCCCACTAAACT 355
Qy 529 ATGATCTCTTATATGCCATCTCCAGCTGGGGCCAGGGGAAAGTATGAATTCACCGAG 588

Db 356 ATGATCTCTTATATGCCATCTCCAAGCTGGGCGAGGAAAGTATGAATTCACGCGAG 415
Qy 589 CGACTGACACCACTCCAGCGTGGTGGTCCATGTATGTGAGAGTATCAAGAGAAGGAGG 648
Db 416 CGACTGACACCACTCCAGCGTGGTGGTCCATGTATGTGAGAGTATCAAGAGAAGGAGG 475
Qy 649 AAGAAGAGGAATGGAAAGATGAGGAGACTTAAGCCAAATAATTCAGACAGCAGAGGC 708
Db 476 AAGAAGAGGAATGGAAAGATGAGGAGACTTAAGCCAAATAATTCAGACAGCAGAGGC 535
Qy 709 CGGAGTACACCGCGATCCACCTCAGCTGAACGCGAGAGGACGCAATTCCAAATC 768
Db 536 CGGAGTACACCGCGATCCACCTCAGCTGAACGCGAGAGGACGCAATTCCAAATC 595
Qy 769 ATACTCAGCGGAGGAATCTTTTACTGTGGAGGTGGTCTCAGCACTTCTCGGAGGTGG 828
Db 596 ATACTCAGCGGAGGA----- 611
Qy 829 CAGCCGAGATCGGGGTGGCAGAAATCCCAAGTTTCTGCTCAGAAGAGATCAAGCGCG 888
Db 612 ----- 611
Qy 889 TGTCCCTCTTCTTAATGCTGCACACCACTTACTGTTTATGGCACTGCTTTCGAGAAATGATCTG 948
Db 612 ----- 611
Qy 949 GCAATCACTGAGTTTGTGGTATCGCACAGGACATTTGGGACTGCTTTCGAGAAACAG 1008
Db 612 ----- 611
Qy 1009 AATAATGATAGTTTGTACTTGTCTTCTGTTTCTGTTAGTGTCTGTTGCAAGCGCAGG 1068
Db 612 ----- 611
Qy 1069 TTGATCAGTGTGCTCAGGAGAGAGCTTCTGTTTCTAAGTGGCCTGCGAGGGCCACTCTC 1128
Db 612 ----- 611
Qy 1129 TACTGTAGGAAGAGTACCAAGAGCGCCTAGTGCAGAGAGGTGTGAAAAACAGCA 1188
Db 612 -----AGGAAGCGCCTAGTGCAGAGAGGTGTGAAAAACAGCA 649
Qy 1189 GCAATGCAATGTGAAATTTGTAGCGTTTCTTCTTCCCTCATGTTCTCATGTTTGTGCA 1248
Db 650 GCAATGCAATGTGAAATTTGTAGCGTTTCTTCTTCCCTCATGTTCTCATGTTTGTGCA 709
Qy 1249 TGTATATTAATGATTTTAAAGACTTAACCTTTGTTTCTGATATAAAGTTTACACCGTCTGTT 1308
Db 710 TGTATATTAATGATTTTAAAGACTTAACCTTTGTTTCTGATATAAAGTTTACACCGTCTGTT 769
Qy 1309 TTTACATCTTTTGGGAGCGCAGGAAAGCGTTTGGAAAAAGTATCACCTTTCCAGATTCT 1368
Db 770 TTTACATCTTTTGGGAGCGCAGGAAAGCGTTTGGAAAAAGTATCACCTTTCCAGATTCT 829
Qy 1369 CGGATCTGCACTCTTTGCAACAGCACTTGTCTTGGCGAACTTCTTCTGGAATGCATTAC 1428
Db 830 CGGATCTGCACTCTTTGCAACAGCACTTGTCTTGGCGAACTTCTTCTGGAATGCATTAC 889
Qy 1429 TCAGCATCCCAACCGTGCAGCGTGAAGTGTGCTTTTGGAAAAAGAGTTGATCTGAAA 1488
Db 890 TCAGCATCCCAACCGTGCAGCGTGAAGTGTGCTTTTGGAAAAAGAGTTGATCTGAAA 949
Qy 1489 TTTCTCTGTAGAATTTAGCTTATACAAATTCAGAGAAATAGCAGTTTCACTGCCAATTTTA 1548
Db 950 TTTCTCTGTAGAATTTAGCTTATACAAATTCAGAGAAATAGCAGTTTCACTGCCAATTTTA 1009
Qy 1549 GTGGGTGAGAAATTTTGTAGTGTGTTTGGGATCGGACCTCAGTTTCTGTTGTTCTTT 1608
Db 1010 GTGGGTGAGAAATTTTGTAGTGTGTTTGGGATCGGACCTCAGTTTCTGTTGTTCTTT 1069
Qy 1609 TATGTGGTGTGTTTCTATACATGATCATAGCCAAATCTTTTGG-AAAATGTTGTTG 1667

Db 770 TTATCATCTTTTGGAGCCAGGAAGCGTTTGGAAAACGTATACCTTTCCAGATTCT 829
Qy 1369 CGGATTCCTGACTCTTTTGCACAGCACTTCTTCCGGAACCTTCTTGGAAATGCATTAC 1428
Db 830 CGGATTCCTGACTCTTTTGCACAGCACTTCTTCCGGAACCTTCTTGGAAATGCATTAC 889
Qy 1429 TCAGCATCCCCAACCGTGCACCGTGTAACTTGTGCTTTTGCAAAAGAGTTGATCTGAAA 1488
Db 890 TCAGCATCCCCAACCGTGCACCGTGTAACTTGTGCTTTTGCAAAAGAGTTGATCTGAAA 949
Qy 1489 TTCTCTGTAGAAATTTAGCTTATACAAATTCAGAGAAATAGCAGTTTCTCAGTCCCAACTTTTA 1548
Db 950 TTCTCTGTAGAAATTTAGCTTATACAAATTCAGAGAAATAGCAGTTTCTCAGTCCCAACTTTTA 1009
Qy 1549 GTGGGTGAGAAATTTTGTAGTTAGGTGTTTGGGATCGGATCGGACTCAGTTTCTGTTTCTTT 1608
Db 1010 GTGGGTGAGAAATTTTGTAGTTAGGTGTTTGGGATCGGACTCAGTTTCTGTTTCTTT 1069
Qy 1609 TATGTGGTGGTTTCTATACATGAATCATAGCCAAAACCTTTTTCGG-AAACTGTTGGTTG 1667
Db 1070 TATGTGGTGGTTTCTATACATGAATCATAGCCAAAACCTTTTTCGG-AAACTGTTGGTTG 1129
Qy 1668 AGATAGTGTGGTTTCTTACCCACGAAGACATCAAGATACACTTGTAAATAAAGCTGATA 1727
Db 1130 AGATAGTGTGGTTTCTTACCCACGAAGACATCAAGATACACTTGTAAATAAAGCTGATA 1189
Qy 1728 GCATATATTCATACCTGTTGTACCTTGGGTGAAAAGATGCGCAGTGGGAGACTAAAGATG 1787
Db 1190 GCATATATTCATACCTGTTGTACCTTGGGTGAAAAGATGCGCAGTGGGAGACTAAAGATG 1249
Qy 1788 TATTAACTTACCTGTGAATCATATGTTGTAGGAAAAGCTGTTCCCATGCTTAACAGGACT 1847
Db 1250 TATTAACTTACCTGTGAATCATATGTTGTAGGAAAAGCTGTTCCCATGCTTAACAGGACT 1309
Qy 1848 TCAATTTCAAAGCATGTCAGTGGATAGATCTGTGGCGGATAGAGGGGATGCACTGC 1907
Db 1310 TGAATTCAAAGCATGTCAGTGGATAGATCTGTGGCGGATAGAGGGGATGCACTGC 1369
Qy 1908 CTTTCCCATTCATTCTCTGATGGAATTTGTATCTAGCTTAAACATTTGTAATTTTCT 1967
Db 1370 CTTTCCCATTCATTCTCTGATGGAATTTGTATCTAGCTTAAACATTTGTAATTTTCT 1429
Qy 1968 AGTTGTATGTATGCTGTGTAATAGGTATATATTTTGGCCTTACAAATACCGTAACA 2027
Db 1430 AGTTGTATGTATGCTGTGTAATAGGTATATATTTTGGCCTTACAAATACCGTAACA 1489
Qy 2028 ATGTTTGTCAFTTTGAAATACTTAATGCCAAGTAAACATGCTTTGGAAATTTGGAA 2087
Db 1490 ATGTTTGTCAFTTTGAAATACTTAATGCCAAGTAAACATGCTTTGGAAATTTGGAA 1549
Qy 2088 GATGGTTTATTTCTTTGAGAAGCAAAATATGTTTGCATTTAAATGCTTTGATTTTCATATC 2147
Db 1550 GATGGTTTATTTCTTTGAGAAGCAAAATATGTTTGCATTTAAATGCTTTGATTTTCATATC 1609
Qy 2148 AAGAAATTTGATTTGAACCTTCTCAACCCCTGTTTACGGTACTTGGTAAGAGGGAGCGGTT 2207
Db 1610 AAGAAATTTGATTTGAACCTTCTCAACCCCTGTTTACGGTACTTGGTAAGAGGGAGCGGTT 1669
Qy 2208 TGGGAGAGACCATTTGCATCGCTGTCCAAAGTGTTCCTTTAAGTGTCTTTAACTGGAGA 2267
Db 1670 TGGGAGAGACCATTTGCATCGCTGTCCAAAGTGTTCCTTTAAGTGTCTTTAACTGGAGA 1729
Qy 2268 GGTAAACCTCAAAATACTTTTAACTGATTTCTATATAAATGGGACAGTATGCTCC 2327
Db 1730 GGTAAACCTCAAAATACTTTTAACTGATTTCTATATAAATGGGACAGTATGCTCC 1789
Qy 2328 TTAC 2331
Db 1790 TTAC 1793

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 14, 2002, 20:00:06 ; Search time 29.8988 Seconds
(without alignments)
2020.659 Million cell updates/sec

Title: US-09-782-953-9

Perfect score: 1041

Sequence: 1 MPAPSMDCVSTLVACVVDV.....SPKPKIQTTRRGLPPSVSN 197

Scoring table:

BLASUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV=xlh

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO spoool/US09782953/runat_11122002_114430_17377/app_query.fasta_1.1173
-DB=Issued_Patents_NA -Qfmt=fastep -SUFFIX=rni -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09782953 @CNC1_1_29 @runat_11122002_114430_17377 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DSELT=7

Database : Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A COMB seq.*
2: /cgn2_6/ptodata/1/ina/5B COMB seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB seq.*
5: /cgn2_6/ptodata/1/ina/6CTUS COMB seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query			DB	ID	Description	
	Score	Match	Length				
1	600	57.6	2174	2	US-08-665-040-1	Sequence 1, Appli	
2	101	9.7	5115	3	US-08-348-518C-3	Sequence 3, Appli	
3	101	9.7	5115	3	US-08-476-509B-3	Sequence 3, Appli	
c	4	89.5	8.6	4411529	4	US-09-103-840A-1	Sequence 1, Appli
	5	89	8.5	3211	2	US-08-574-959A-8	Sequence 8, Appli
	6	89	8.5	3211	4	US-09-357-014-8	Sequence 8, Appli
	7	89	8.5	3901	2	US-08-574-959A-6	Sequence 6, Appli
8	89	8.5	3901	4	US-09-357-014-6	Sequence 6, Appli	
9	83.5	8.0	1869	3	US-09-372-668-3	Sequence 3, Appli	
c	10	82.5	7.9	3483	3	US-08-808-599A-23	Sequence 23, Appli
11	82.5	7.9	4411529	4	US-09-103-840A-1	Sequence 1, Appli	
12	82	7.9	4112	1	US-08-340-203A-2	Sequence 2, Appli	

13	82	7.9	4112	2	US-08-452-567-2	Sequence 2, Appli
14	82	7.9	4112	2	US-08-452-427-2	Sequence 2, Appli
15	82	7.9	4112	3	US-09-085-407-2	Sequence 2, Appli
16	82	7.9	4616	1	US-08-340-203A-1	Sequence 1, Appli
17	82	7.9	4616	2	US-08-452-567-1	Sequence 1, Appli
18	82	7.9	4616	2	US-08-452-427-1	Sequence 1, Appli
19	82	7.9	4616	3	US-09-085-407-1	Sequence 1, Appli
20	80	7.7	3089	1	US-07-921-796-7	Sequence 7, Appli
21	79.5	7.6	5410	4	US-09-221-017B-70	Sequence 20, Appli
22	79	7.6	169998	4	US-09-676-610B-24	Sequence 74, Appli
23	78	7.5	3360	1	US-08-408-093-5	Sequence 5, Appli
24	78	7.5	3360	1	US-08-408-420A-5	Sequence 5, Appli
25	78	7.5	3360	1	US-08-714-901-5	Sequence 5, Appli
26	78	7.5	3360	3	US-08-040-741-5	Sequence 2, Appli
27	78	7.5	3468	1	US-07-951-715A-2	Sequence 2, Appli
28	78	7.5	3468	1	US-07-951-715A-4	Sequence 4, Appli
29	78	7.5	3468	2	US-08-459-448A-2	Sequence 4, Appli
30	78	7.5	3468	2	US-08-459-448A-4	Sequence 4, Appli
31	78	7.5	3468	3	US-08-459-595A-2	Sequence 2, Appli
32	78	7.5	3468	3	US-08-459-595A-4	Sequence 2, Appli
33	78	7.5	3468	3	US-08-459-504B-2	Sequence 2, Appli
34	78	7.5	3468	3	US-08-459-504B-4	Sequence 2, Appli
35	78	7.5	3468	3	US-08-459-444-2	Sequence 2, Appli
36	78	7.5	3468	3	US-08-459-444-4	Sequence 2, Appli
37	78	7.5	3468	3	US-09-053-549-3	Sequence 3, Appli
38	78	7.5	3468	3	US-09-053-549-5	Sequence 5, Appli
39	78	7.5	3468	4	US-09-547-422-2	Sequence 2, Appli
40	78	7.5	3468	4	US-09-547-422-4	Sequence 4, Appli
C	41	78	7.5	80161	3	US-09-036-987A-1
C	42	78	7.5	80161	4	US-09-370-700-1
C	43	77.5	7.4	2524	1	US-08-317-522A-1
C	44	77.5	7.4	2524	1	US-08-439-818A-1
C	45	77.5	7.4	2524	2	US-08-751-965-1

ALIGNMENTS

.RESULT 1
US-08-665-040-1
; Sequence 1, Application US/08665040
; Patent No. 5869318
; GENERAL INFORMATION:
; APPLICANT: ESTIVILL PALLEJA, XAVIER
; APPLICANT: FUENTES, JUAN JOSE
; APPLICANT: PRITCHARD, MELANIE
; TITLE OF INVENTION: A NEW GENE SEQUENCE OF THE
; TITLE OF INVENTION: DOWN SYNDROM CRITICAL REGION OF HUMAN CHROMOSOME 21,
; TITLE OF INVENTION: IDENTIFIED BY A NEW "ALU-SPLICING PCR" TECHNIQUE,
; TITLE OF INVENTION: CODING FOR A PROLINE-RICH PROTEIN (DSCR1) HIGHLY
; TITLE OF INVENTION: EXPRESSED IN FOETEL BRAIN AND IN HEART AND METHOD
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LADAS & PARRY
; STREET: 26 WEST 61ST STREET
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10023
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORD PERFECT 5.1 FOR DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/665,040
; FILING DATE: JUNE 7, 1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: ES P9501140
; FILING DATE: JUNE 7, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: JANET I. CORD

Feature	Brain	Brain
REGISTRATION NUMBER:	33,778	
REFERENCE/DOCKET NUMBER:	U010815-9	
TELECOMMUNICATION INFORMATION:		
TELEPHONE:	(212) 708-1800	
TELEFAX:	(212) 246-8959	
INFORMATION FOR SEQ ID NO:	1:	
SEQUENCE CHARACTERISTICS:		
LENGTH:	2174 base pairs	
TYPE:	nucleic acid	
STRANDEDNESS:	double	
TOPOLOGY:	circular	
MOLECULE TYPE:	cdna for mRNA	
HYPOTHETICAL:	No	
ANTI-SENSE:	No	
ORIGINAL SOURCE:	human	
ORGANISM:	Homo sapiens	
DEVELOPMENTAL STAGE:	foetal	
TISSUE TYPE:	Brain	
IMMEDIATE SOURCE:	gene library of cDNA	
LIBRARY:	gene library of cDNA from foetal	
CLONE:	BC-17.8-1 and BC-17.8-2	
POSITION IN GENOME:		
CHROMOSOME/SEGMENT:	chromosome 21/YAC 72H9	
MAP POSITION:	21q22.1-q22.2	
FEATURE:		
NAME/KEY:	CDNA for mRNA, BC-17.8	
LOCATION:	1..2174	
FEATURE:		
NAME/KEY:	untranslated 5'	
LOCATION:	1..48	
FEATURE:		
NAME/KEY:	coding sequence	
LOCATION:	49..560	
OTHER INFORMATION:	Down Syndrome critical	
FEATURE:		
NAME/KEY:	DSCL1	
LOCATION:	1..171 PEPTIDES	
IDENTIFICATION METHOD:	translation of the	
OTHER INFORMATION:		
OTHER INFORMATION:	- deduced protein	
OTHER INFORMATION:	- proline-rich protein domains	
OTHER INFORMATION:	- glutamic acid-rich protein domains	
OTHER INFORMATION:	- leucine/phenylalanine-rich protein domains	
FEATURE:		
NAME/KEY:	untranslated 3'	
LOCATION:	564..2174	
FEATURE:		
NAME/KEY:	2 poly (A)	
LOCATION:	1541..1546 AND 2132..2137	
US-08-665-040-1		
Alignment Scores:		
Pred. No.:	1,74e-58	2174
Score:	600.00	115
Percent Similarity:	79.77%	Matches: 23
Best Local Similarity:	66.47%	Mismatches: 27
Query Match:	57.64%	Indels: 8
DB:	2	Gaps: 2
US-09-782-953-9 (1-197) x US-08-665-040-1 (1-2174)		
QY	25	25
DB	43	43
QY	45	45
DB	103	103
QY	65	65
DB	163	163

```

Oy 85 PheaaGlnValaGlnThrProGluThrAspGlyAspLeuHisLeuAlaProProGln 104
Db 223 TTtGCTCAGACCTTACACATAGAAAGCTCA-----CACCTGGCTCCGCCAAT 270
Oy 105 ProAlaLysGlnPheLeuLeuIleSerProProSerSerProProValGlyTyrLysProIle 124
Db 271 CCAGCAAGACGATTTCTGATCTCTCCCTCCGCCCTCTCCGCAAGTGGAGTGAACCACTG 330
Oy 125 SerAspAlaThrProValLeuAsnTyrAspLeuLeuTyrAlaValAlaLysGlyPro 144
Db 331 GAAGATGGACCCCAAGCATAACTATGATCTCTTATATGTCACATCTCCAACTCGGGCCCA 390
Oy 145 GlyGluLysTyrGluLeuHisAlaGlyThrGluSerThrProSerValValHisVal 164
Db 391 GGGGAAAGATGATGATCAAGAGAAGAGAGAAAGAAATGAAAGATGAGAGAGACTTAG 450
Oy 165 CysAspSerAspMetGluGluGluGluLysAspProLysThrSer-----ProLys 180
Db 451 TGTGAGACTGATCAAGAGAAGAGAGAAAGAAATGAAAGATGAGAGAGACTTAG 510
Oy 181 ProLysIleIleGlnThrArgArgProGlyLeuProPro 193
Db 511 CCAAAATATATCCAGACGAGGCCGCGATACAGCGG 549

RESULT 2
US-08-348-518C-3
; Sequence 3, Application US/08348518C
; Patent No. 6022740
; GENERAL INFORMATION:
; APPLICANT: SUDOL, MARIUS
; APPLICANT: PEER, BORK
; APPLICANT: HENRY, CHEN
; TITLE OF INVENTION: A SH2 DOMAIN ASSOCIATED PROTEIN, A
; TITLE OF INVENTION: SIGNALING DOMAIN THEREIN, NUCLEIC ACIDS ENCODING THE
; TITLE OF INVENTION: PROTEIN AND THE DOMAIN, AND DIAGNOSTIC AND THERAPEUTIC USES
; TITLE OF INVENTION: THEROP
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Klauder & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/348,518C
; FILING DATE: 01-DEC-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Bsq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 515 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Human
; IMMEDIATE SOURCE:

```



```

; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37RV
US-09-103-840A-1

Alignment Scores:
Pred. No.: 5.08e+04 Length: 4411529
Score: 89.50 Matches: 53
Percent Similarity: 35.56% Conservative: 27
Best Local Similarity: 23.56% Mismatches: 85
Query Match: 8.60% Indels: 60
DB: 4 Gaps: 10

US-09-782-953-9 (1-197) x US-09-103-840A-1 (1-4411529)
Oy 8 CyaaPvaPvaLSeThrLeuValAla-----Cys---Val 17
Db 4094824 TCCCGCGCTGCGAGCGCTGTGCTGCGCTGCGCTGCGAGCAGCCGCACTGCGAGGTG 4094765
Oy 18 ValaPvaLgluValPheThraNgluValLyvgLylPhegluLylLeuPheArg 37
Db 4094764 GTGAGCCTCGAGCGTGTGTCACCGTGAAGGTGCTGCGTTCGCGGTGTGCGCAGC 4094705
Oy 38 ThrTyraPgluCyValThrPhegluLeu----- 47
Db 4094704 GCCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4094645
Oy 48 -----PheLysSerPheArgArgValArgIleAsnPhe-----SerHisPro 61
Db 4094644 TCTATCGGCAAGCGAACTATCGGCAAGCGAACTAGTGTGCGGTACGGGAGAACCA 4094585
Oy 62 LyvSerAlaAlaArgAlaArgIleLeuValThrgLInPheArgLylLysLysLeu 81
Db 4094584 ACTACGGCGCAGCAGCTCGCGGTGTGATCCCGGAGAACAG--CGCGCGCGCGCGCAG 4094528
Oy 82 LylLeu-----TyrPheAlaGlnValGlnThr-ProGluThrAspG1 95
Db 4094527 CGGCTACCTCAACCGCGTATGCGCTCTGCGCGCGAGCAGCGCGCGCG----- 4094479
Oy 95 YaaPryLysuHlLysValaProGlnProAlaLysGlnPheLeuLlSerProProse 115
Db 4094478 -----GTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4094435
Oy 115 rSerProProValGlyTyrLysProLlSerAspAlaThrProValLeuAsnTyrAspLe 135
Db 4094434 GTTACACCGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4094392
Oy 135 uLeuTyraLAlaLysLeuGlyProGlyLylLysTyrGlyLysLeuHlAlaGlyThrG1 155
Db 4094391 -----CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4094339
Oy 155 userThrProSerValValIleValIleCysAspSer-----AspMetG1 170
Db 4094338 CAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4094279
Oy 170 uGluGluGluAspProLysLleIleGlnThrArgArgProG1 190
Db 4094278 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4094231
Oy 190 YleuProProser 194
Db 4094230 GTTGC CGCGCGCG 4094218

RESULT 5
US-08-574-959A-8
; Sequence 8, Application US/08574959A
; Patent No. 5962224
; GENERAL INFORMATION:
; APPLICANT: Jaekyoon Shin, Insil Jyoung, Ratna K. Vadlamudi
; TITLE OF INVENTION: p62 POLYPEPTIDES, RELATED POLYPEPTIDES
; TITLE OF INVENTION: AND USES THEREFOR

```

```

; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/574,959A
; FILING DATE: 19-DEC-95
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: DFN-008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3211 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 439..3157
; US-08-574-959A-8

Alignment Scores:
Pred. No.: 2.49 Length: 3211
Score: 89.00 Matches: 29
Percent Similarity: 42.86% Conservative: 13
Best Local Similarity: 29.59% Mismatches: 44
Query Match: 8.55% Indels: 12
DB: 2 Gaps: 5

US-09-782-953-9 (1-197) x US-08-574-959A-8 (1-3211)
Oy 101 AlaProGlnProAlaLysGlnPheLeuLlSerProProserProProValGly 120
Db 2689 GCACACACACCCCAACCTGCTCCAGCTGCTGCTCCCTCCCTGAGTCTCCCA--AAG 2745
Oy 121 TrpLysProLlSerAspAlaThrProValLeuAsnTyrAspLeuLylLysLysLeu 140
Db 2746 GTGACGCGAACCCGCAACCCGCAACCC-----GGGCTGCTTTTGGAACTGGAG 2793
Oy 141 LylLeuGlyProGlyGluLysTyrGlyLysLeuHlAlaGlyThrGlnSerThrProSerVal 160
Db 2794 GAG-----CCAGGAGCAGAGAGAG-----CGTGAAGCTGACACAGCTCCACCTCG 2841
Oy 161 ValValHlLysValCysAspSerAspMetGlu-----GluGluGluAspProLysThr 177
Db 2842 GCGCTGAAGCGCTCCCTCCAGGAGAGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2901
Oy 178 SerProLysProLysLleIleGlnThrArgArgProGlyLeuProProSerVal 195
Db 2902 GGGCCCGCTCCCGCAGAGCTGTGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2955

RESULT 6
US-09-357-014-8
; Sequence 8, Application US/09357014
; Patent No. 6291645
; GENERAL INFORMATION:
; APPLICANT: Jaekyoon Shin, Insil Jyoung, Ratna K. Vadlamudi
; TITLE OF INVENTION: p62 POLYPEPTIDES, RELATED POLYPEPTIDES

```



```

Db 543 CAGGTGACCCCTGGAGAGCCCGCCATGATGAGCTTACACACACCCACCGCCACT 602
Qy 181 ProLysillelleGlnThrArgProGlyLeuProSerValSer 196
Db 603 GGGGTCTTCTCCCTCAAGGCCCGGCTTGGCTCCCACTGGGATCAAC 650

RESULT 10
US-08-808-599A-23/c
; Sequence 23, Application US/08080599A
; Patent No. 6111089
; GENERAL INFORMATION:
; APPLICANT: Fukuda, Michiko N.
; TITLE OF INVENTION: Trophinin, Trophinin-Assisting
; TITLE OF INVENTION: Proteins and Methods to Inhibit Implantation
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/808,599A
; FILING DATE: 28-FEB-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/317,522
; FILING DATE: 04-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/439,818
; FILING DATE: 12-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LA 2256
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3483 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3481
US-08-808-599A-23

Alignment Scores:
Pred. No.: 15 2 Length: 3483
Score: 82.50 Matches: 31
Percent Similarity: 41.75% Conservative: 12
Best Local Similarity: 30.10% Mismatches: 49
Query Match: 7.93% Indels: 11
DB: 3 Gaps: 4

US-09-782-953-9 (1-197) x US-08-808-599A-23 (1-3483)
Qy 100 LeuAlaProGlnProLysGlnPheLeuIleSerProProSerSerProProVal 119
Db 3164 CTTGGTCCACCGAAGCCAGGCTCTCTCTGGTGGTCCACCCAGCCACCGATG 3111
Qy 120 GlyTyrIleProIleSerAspAlaThrProValLeuAsnTyrAspLeuTyrAlaVal 139
Db 3110 ---CTTGGTCCACTCGAGAGGCCAGCAGCAGTGTGGTTCACCACTAAAGCCACCAATA 3054

```

```

QY 140 AlaLysLeuGlyProGlyGluLysThrGluLeuHisAlaGlyThrGluSerThrProSer 159
Db 3053 ---GAAGCTGGTCCATTACTGAAGCTGGAGCCAGTGTGGGTTCCACGATAAGCCATTG 2997
QY 160 valvalvalHisvalCysAspSeraspMet-----GluGluGlu 172
Db 2996 CTGTGTGTTGGAAACCACTGCCAAGCAATGATGGTATTTCAGTCTCTGTTGAAGCTGGCA 2937
QY 173 GluAspProLysThrSerProLysProLysIlelleGlnThrArgArgProGlyLeuPro 192
Db 2936 TTAGGCCCATTCGCAAGCAATGCCATTGTGGAGCTTACACTAAAGCCTGCACCAAGTG 2877
QY 193 ProSerVal 195
Db 2876 CCAAGTGTG 2868

RESULT 11
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103.840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Alignment Scores:
Pred. No.: 2,71e+05 Length: 4411529
Score: 82.50 Matches: 31
Percent Similarity: 43.48% Conservative: 9
Best Local Similarity: 33.70% Mismatches: 32
Query Match: 7.93% Indels: 20
DB: 4 Gaps: 5

US-09-782-953-9 (1-197) x US-09-103-840A-1 (1-4411529)
QY 50 SerPheArgArgVal-----ArgIleAsnPheSerHisProLysSerAlaAlaArgAla 67
Db 3667235 TCGTCTCTGGAGGTCAACACGCGCTTCAGGTAGACACCGGTCCAGGAGAAACCGCG 36672394
QY 68 ArgIleGluLeuHisGluThrGlnPheArg-----GlyLysLysLeuLysLeuTyr 84
Db 3667295 GGCATGACATTGGTGTGTGACGAATTCGGCATCGCAACGCGCAAAAGCTGGACATC--- 3667351
QY 85 PheAlaGlnValGlnThrProGlu-----Thrasp 94
Db 3667352 ---ACGAGGATCCCAACCGCGCGGCAAGCCATCGAATTCGGATCAACGCGGAGGAC 3667408
QY 95 GlyAspLysLeuHisLeuAlaProGlnProAlaLysGlnPheLeuLysProPro 114
Db 3667409 GCGGGGGTAACTTCTTACCGCGCGCGGGCCGGTGACAAGATTCTC-----CACCCGCGC 3667462
QY 115 SerSerProProVal-GlyTrpLysProLysSer 125
Db 3667463 TCCGGGCGCGGTGTGCGGTGGACTCCGGTGTGCG 3667496

RESULT 12
US-08-340-203A-2
; Sequence 2, Application US/08340203A
; Patent No. 5756668

```

GENERAL INFORMATION:
APPLICANT: Baylin, Stephen B.
APPLICANT: Maies, Michele M.
TITLE OF INVENTION: NOVEL TUMOR SUPPRESSOR GENE, HIC-1
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: California
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/340,203A
FILING DATE: 15-NOV-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Ph.D., Lisa A.
REGISTRATION NUMBER: P-38,347
REFERENCE/DOCKET NUMBER: 07265/039001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 455-5100
TELEFAX: (619) 455-5110
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 4112 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
IMMEDIATE SOURCE:
CLONE: HIC-1 coding polynucleotide
FEATURE:
NAME/KEY: CDS
LOCATION: 1086..2726
US-08-340-203A-2
Alignment Scores:
Pred. No.: 21.8 Length: 4112
Score: 82.00 Matches: 41
Percent Similarity: 33.96% Conservative: 13
Best Local Similarity: 25.79% Mismatches: 41
Query Match: 7.88% Indels: 64
Gaps: 9
US-09-782-953-9 (1-197) x US-08-340-203A-2 (1-4112)
QY 93 TThApGlyAspLyLeuHIsLeuAlaProProGlnProAlaLeuGlnPheLeuIleSer 112
DB 2975 ACCAAAGGAGACCCAGGCCCTC---CCGCCCTTCTGTGTGTCGTC-----CGG 3022
QY 113 ProProSerSerProPro----- 118
DB 3023 CCCCCTCCCCCGGCTCCGCCCTCTTAAAGGGAGAGGGGTGTCACTGCGGGCACTC 3082
QY 119 ValGlyTTrpLyAspProIleSerAspAlaThrProValLeuLeuAntyAspLeuLeuTyraIa 138
DB 3083 CTAGACCTACTCTCCGGCCCTTGGACACACCATCTCTCACTGTAATCTCCCGCGTGGG 3142
QY 139 Val-----AlaLyLeuGlyProGlyGlyLeuTyTrpIu----- 149
DB 3143 TCGAGAGCTCGGCGAGAGTGGAGAGGAGGAGCTAGCGCGCGGAGAGCCCGCCGC 3202
QY 150 -----LeuHIsAlaGlyThrGluSerThrProSerValValHis----- 163
DB 3203 ACCCGCGCTCCACCCACCCCGGAGCTGATATATGTAAGTTCCATTTGCAACAAGTGG 3262
QY 164 -----ValCyAspSer 167

DB 3263 CACTAGCCCAAGGCCAACCTTCTCTCTCAATGACCAAGAGGGGAGGATTCGAGACTG 3322
QY 168 AspMetGluGluGluAspProLyThr---SerProLyAspProLyIle----- 183
DB 3323 -----GAGGCGAAGAGCTTACACAGAGTCTCCCACTCCCGGAGGCCCTCCTT 3373
QY 184 -----IleGlnThrArgArgProGlyLeuProProSer 194
DB 3374 CCCTTCCCTCGCGCCCGGAGCAATATTATTGATGAGCGCCCGCGCCCAATCC 3430
RESULT 13
US-08-452-567-2
Sequence 2, Application US/08452567
Patent No. 5846712
GENERAL INFORMATION:
APPLICANT: Baylin, Stephen B.
APPLICANT: Maies, Michele M.
TITLE OF INVENTION: NOVEL TUMOR SUPPRESSOR GENE, HIC-1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: California
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/452,567
FILING DATE: 25-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Ph.D., Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07265/039003 (Div PD3664)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 678-5070
TELEFAX: (619) 678-5099
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 4112 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
IMMEDIATE SOURCE:
CLONE: HIC-1 coding polynucleotide
FEATURE:
NAME/KEY: CDS
LOCATION: 1086..2726
US-08-452-567-2
Alignment Scores:
Pred. No.: 21.8 Length: 4112
Score: 82.00 Matches: 41
Percent Similarity: 33.96% Conservative: 13
Best Local Similarity: 25.79% Mismatches: 41
Query Match: 7.88% Indels: 64
Gaps: 9
US-09-782-953-9 (1-197) x US-08-452-567-2 (1-4112)
QY 93 TThApGlyAspLyLeuHIsLeuAlaProProGlnProAlaLeuGlnPheLeuIleSer 112
DB 2975 ACCAAAGGAGACCCAGGCCCTC---CCGCCCTTCTGTGTGTCGTC-----CGG 3022
QY 113 ProProSerSerProPro----- 118

Db 3023 CCCCCCTCCCGCTCCGGCTGCTCTTAGAGGGGAGGGGTGTCACTGTTCGGGGCACTC 3082
QY 119 ValGlyTrpLysProIleSerAspAlaThrProValLeuAsnTyrAspLeuTyrAla 138
Db 3083 CTAGCCCTACCTCCGGCCCTTCGGACACACCCATCTCTCACTGTGAATCTCCCGCTGGG 3142
QY 139 Val-----AlaValLeuGlyProGlyGluLysTyrGlu----- 149
Db 3143 TCGGAGCGTTCGGGAGAGTGGGGAGTGGGGAGGAGTGGAGCGCGGAGGCGCCCGC 3202
QY 150 -----LeuHisAlaGlyThrGluSerThrProSerValValValHis----- 163
Db 3203 ACCCGCGCTCCACCCACCGGAGTGTATGTGAAGTTCCTCATTTTGCAAGTGG 3262
QY 164 -----ValCysAspSer 167
Db 3263 CACTAGCCAGCGCCACCCCTTCTCTCAGTCACCAAGGCGGGAGTCTCGAGTGG 3322
QY 168 AspMetGluGluGluAspProLysThr---SerProLysProLysIle----- 183
Db 3323 -----GAAGCGAAGAGCCTACACAGGCTCTCCACTCCCGGGGTGCCCTCCCTT 3373
QY 184 -----IleGlnThrArgArgProGlyLeuProProSer 194
Db 3374 CCCTTCCCTCGCGCGCCGACCATATTTATTGTCATGCGCGCGCGCGCCCATCC 3430
RESULT 14
US-08-452-427-2
; Sequence 2, Application US/08452427
; Patent No. 5922590
; GENERAL INFORMATION:
; APPLICANT: Baylin, Stephen B.
; APPLICANT: Wales, Michele M.
; TITLE OF INVENTION: NOVEL TUMOR SUPPRESSOR GENE, HIC-1
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: California
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 26-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/340,203
; FILING DATE: 15-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Ph.D., Lisa A.
; REGISTRATION NUMBER: P-38,347
; REFERENCE/DOCKET NUMBER: 07265/039001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 455-5100
; TELEFAX: (619) 455-5110
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4112 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1086...2726

US-08-452-427-2

Alignment Scores:
Pred. No.: 21.8 Length: 4112
Score: 82.00 Matches: 41
Percent Similarity: 33.96% Conservative: 13
Best Local Similarity: 25.79% Mismatches: 41
Query Match: 7.88% Indels: 64
DB: 2 Gaps: 9

US-09-782-953-9 (1-197) x US-08-452-427-2 (1-4112)

QY 93 ThrAspGlyAspLysLeuHisLeuAlaProGlnProAlaLysGlnPheLeuIleSer 112
Db 2975 ACCAAGAGAGACCCAGGCCCTC-----CCGCTTCTCTGTGTGTTCTG-----CGG 3022
QY 113 ProProSerSerProPro----- 118
Db 3023 CCCCCCTCCCGCTCCGGCTGCTCTTAGAGGGGAGGGGTGTCACTGTTCGGGGCACTC 3082
QY 119 ValGlyTrpLysProIleSerAspAlaThrProValLeuAsnTyrAspLeuTyrAla 138
Db 3083 CTAGCCCTACCTCCGGCCCTTCGGACACACCCATCTCTCACTGTGAATCTCCCGCTGGG 3142
QY 139 Val-----AlaValLeuGlyProGlyGluLysTyrGlu----- 149
Db 3143 TCGGAGCGTTCGGGAGAGTGGGGAGTGGGGAGGAGTGGAGCGCGGAGGCGCCCGC 3202
QY 150 -----LeuHisAlaGlyThrGluSerThrProSerValValValHis----- 163
Db 3203 ACCCGCGCTCCACCCACCGGAGTGTATGTGAAGTTCCTCATTTTGCAAGTGG 3262
QY 164 -----ValCysAspSer 167
Db 3263 CACTAGCCAGCGCCACCCCTTCTCTCAGTCACCAAGGCGGGAGTCTCGAGTGG 3322
QY 168 AspMetGluGluGluAspProLysThr---SerProLysProLysIle----- 183
Db 3323 -----GAAGCGAAGAGCCTACACAGGCTCTCCACTCCCGGGGTGCCCTCCCTT 3373
QY 184 -----IleGlnThrArgArgProGlyLeuProProSer 194
Db 3374 CCCTTCCCTCGCGCGCCGACCATATTTATTGTCATGCGCGCGCGCGCCCATCC 3430
RESULT 15
US-09-085-407-2
; Sequence 2, Application US/09085407
; Patent No. 6103877
; GENERAL INFORMATION:
; APPLICANT: Baylin, Stephen B.
; APPLICANT: Wales, Michele M.
; TITLE OF INVENTION: NOVEL TUMOR SUPPRESSOR GENE, HIC-1
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: California
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/085,407
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/340,203
; FILING DATE: 15-NOV-1994
; ATTORNEY/AGENT INFORMATION:

NAME: Haile, Ph.D., Lisa A.
REGISTRATION NUMBER: P-38,347
REFERENCE/DOCKET NUMBER: 07265/039001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 455-5100
TELEFAX: (619) 455-5110
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 4112 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
IMMEDIATE SOURCE:
CLONE: HIC-1 coding polynucleotide
FEATURE:
NAME/KEY: CDS
LOCATION: 1086..2726
US-09-085-407-2

Alignment Scores:
Pred. No.: 21.8 Length: 4112
Score: 82.00 Matches: 41
Percent Similarity: 33.96% Conservative: 13
Best Local Similarity: 25.79% Mismatches: 41
Query Match: 7.88% Indels: 64
Gaps: 9
DB:

US-09-782-953-9 (1-197) X US-09-085-407-2 (1-4112)

QY 93 ThrAspGlyAspIysLeuHisLeuAlaProProGlnProAlaIysGlnPheLeuIleSer 112
DB 2975 ACCAAGAGAGACCCCAAGGCCCTC---CGGCTCTTCCTGTGTGCTGCT-----CGG 3022

QY 113 ProProSerSerProPro----- 118
DB 3023 CCCCCTCCCCCGGCTCGCGCTCTTAGAGGGAGGGGTGTCACTGTGGGGCACTC 3082

QY 119 ValGlyTyrIlyProIleSerAspAlaThrProValIleuAsnTyrAspLeuLeuTyrAla 138
DB 3083 CTAGCCCTACCTCCGGCCCTTGCGACCACTTCTCACTGTGAATCTCCCGGCTGGG 3142

QY 139 Val-----AlaIysLeuGlyProGlyGluIlyTyrGlu----- 149
DB 3143 TCGGAGCGTGGGAGAGCTTGGGAGTGGGAGGAGCTGAGCCGCGCGAGGCCCGCC 3202

QY 150 -----LeuHisAlaGlyThrIleSerThrProSerValValValHis----- 163
DB 3203 ACCCCGCTCCACCCACCCCGGAGTGAATGTGAAGTTCTCATTTGCAACAGTGG 3262

QY 164 -----ValCysAspSer 167
DB 3263 CACTAGCCCAAGGCCAACCTTCTCTCTCAGTACCAAGGCGGGAGTTCTGAGTGG 3322

QY 168 AspMetGluGluGluAspProIlyThr---SerProIlyProIlySile----- 183
DB 3323 -----GAAAGCGAAGAGCTTACCACTGCTCCACTCCCGGCTGCTCCCTT 3373

QY 184 -----IleGlnThrArgProGlyLeuProProSer 194
DB 3374 CCTTCCCTGGGCCCCGAGACATATTATTGCAATGCGCCCCGGGCCCCCAATCC 3430

Search completed: December 14, 2002, 22:55:07
Job time : 1511.9 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2002, 17:10:31 ; Search time 20.3119 Seconds
(without alignment)
11623.258 Million cell updates/sec

Title: US-09-782-953-5

Perfect score: 597

Sequence: 1 atggattttgggacttag.....caccgatccacttagctga 597

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 355320 seqs, 197730502 residues

Total number of hits satisfying chosen parameters: 710640

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	597	100.0	597	10	US-09-782-953-5
2	515.8	86.4	597	10	US-09-782-953-2
3	504.4	84.5	599	10	US-09-782-953-1
4	470.6	78.8	2331	10	US-09-782-953-11
5	470.6	78.8	2355	12	US-10-044-090-255
6	408.2	68.4	2212	10	US-09-782-953-14
7	399.6	66.9	2173	10	US-09-880-107-3340
8	395.4	66.2	2358	10	US-09-925-302-347
9	214.2	35.9	3184	10	US-09-954-456-497
10	214.2	35.9	3184	10	US-09-782-953-17
11	201.4	33.7	594	10	US-09-782-953-8
12	191.2	32.0	720	10	US-09-782-953-23
13	191.2	32.0	828	10	US-09-782-953-20
14	144.6	24.2	412	10	US-09-864-761-10388
15	144.6	24.2	446	10	US-09-864-761-2064
16	123	20.6	486	10	US-09-864-761-809
17	91.8	15.4	111	10	US-09-864-761-27019
18	84.4	14.1	123	10	US-09-864-761-17592
19	81.4	13.6	365	10	US-09-728-445-736

C	20	70.6	11.8	85	10	US-09-864-761-18808	Sequence 18808, A
	21	37.8	6.3	2853	10	US-09-965-631-3	Sequence 3, Appli
	22	37.8	6.3	3446	10	US-09-965-631-7	Sequence 7, Appli
	23	36.2	6.1	5739	10	US-09-960-253-142	Sequence 142, App
	24	36.2	6.1	6417	10	US-09-962-436-288	Sequence 288, App
	25	35.6	6.0	490	10	US-09-880-107-2261	Sequence 2261, Ap
	26	34.4	5.8	497	10	US-09-822-263-35	Sequence 35, Appl
	27	34.4	5.8	4363	10	US-09-764-864-241	Sequence 17, Appl
	28	34	5.7	345	10	US-09-822-263-17	Sequence 241, App
	29	34	5.7	387	10	US-09-920-300A-395	Sequence 395, App
	30	34	5.7	387	12	US-10-033-528-392	Sequence 392, App
	31	34	5.7	438	9	US-09-736-457-1028	Sequence 1028, Ap
	32	34	5.7	438	9	US-09-902-941-1028	Sequence 1028, Ap
	33	34	5.7	497	10	US-09-822-263-33	Sequence 33, Appl
	34	34	5.7	1147	10	US-09-880-107-2311	Sequence 2311, Ap
	35	33.8	5.7	156	10	US-09-864-761-20014	Sequence 20014, A
	36	33.8	5.7	415	10	US-09-864-761-3236	Sequence 3236, Ap
	37	33.6	5.6	1752	9	US-09-887-552A-1	Sequence 1, Appli
	38	33.4	5.6	565	10	US-09-962-436-116	Sequence 116, App
	39	33.4	5.6	565	10	US-09-934-456-1735	Sequence 1735, Ap
	40	33.4	5.6	575	10	US-09-864-761-8864	Sequence 8864, Ap
	41	33.4	5.6	20247	10	US-09-764-877-2680	Sequence 2680, Ap
	42	33.2	5.6	299	10	US-09-864-761-21553	Sequence 21553, A
	43	33.2	5.6	470	10	US-09-864-761-2442	Sequence 2442, Ap
	44	33.2	5.6	1875	9	US-10-001-835-56	Sequence 56, Appl
	45	33.2	5.6	2289	10	US-09-850-964-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1

US-09-782-953-5

Sequence 5, Application US/09782953

Patent No. US20020150953A1

GENERAL INFORMATION:

APPLICANT: WILLIAMS, R. SANDERS

APPLICANT: ROTHERMEL, BEVERLY

TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
CALCINEURIN INTERACTING PROTEIN (MCIP)

FILE REFERENCE: UTSD:674P21

CURRENT APPLICATION NUMBER: US/09/782,953

CURRENT FILING DATE: 2001-02-13

PRIOR APPLICATION NUMBER: 60/216,601

PRIOR FILING DATE: 2000-07-07

NUMBER OF SEQ ID NOS: 27

SOFTWARE: Patentin ver. 2.1

SEQ ID NO 5

LENGTH: 597

TYPE: DNA

ORGANISM: Mus musculus

FEATURE:

NAME/KEY: CDS

LOCATION: (1)...(594)

US-09-782-953-5

Query Match

Best Local Similarity 100.0%; Score 597; DB 10; Length 597;

Matches 597; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGATTTAGGACTTTAGCTACAATTTTAGCTCCCTGATTTGCTGTGGCAACGAT 60

Db 1 ATGGATTTAGGACTTTAGCTACAATTTTAGCTCCCTGATTTGCTGTGGCAACGAT 60

Qy 61 GATCTCTTACGGAAGTGAGCCAGGCGCAAAATTTGAATCCCTCTTTCAGAACATATGAC 120

Db 61 GATCTCTTACGGAAGTGAGCCAGGCGCAAAATTTGAATCCCTCTTTCAGAACATATGAC 120

Qy 121 AAGGACACCACTTCCAGTAGTATTTTAAGAGCTTCAAACGTGTCGGATAAACTTCAGCAAC 180

Db 121 AAGGACACCACTTCCAGTAGTATTTTAAGAGCTTCAAACGTGTCGGATAAACTTCAGCAAC 180

Qy 181 CCCTTATCTGAGCCGATGCCGCTCGCGCTGCAACAGCCGAGTTCTCTGGGGAAGGAA 240

```

|||||
Db 181 CCCATATGAGCCGATCCAGGCTGGGCTGCACAAAGCCGAGTTCCTGGGAGAA 240
Qy 241 ATGAAGTTGATTTTGGCTCAGACTTTACATAGAGAGTTCAACCTGGCTCCGCCAAT 300
Db 241 ATGAAGTTGATTTTGGCTCAGACTTTACATAGAGAGTTCAACCTGGCTCCGCCAAT 300
Qy 301 CCCGACAAAGCTTCTCATCTCCCTCCGCTCTCTCCGTTGGTGAACAAAGTA 360
Db 301 CCCGACAAAGCTTCTCATCTCCCTCCGCTCTCTCCGTTGGTGAACAAAGTA 360
Qy 361 GAGATGCGACCCCGCTGATAAATTACGATCTTTATATGCGATCTCCAGCTGGGCCA 420
Db 361 GAGATGCGACCCCGCTGATAAATTACGATCTTTATATGCGATCTCCAGCTGGGCCA 420
Qy 421 GGAGAGAGTATGAATGCTGATGACGACGACAGCCCACTCCAGTGTGTGTCACGCTG 480
Db 421 GGAGAGAGTATGAATGCTGATGACGACGACAGCCCACTCCAGTGTGTGTCACGCTG 480
Qy 481 TGTGAGAGTACCAAGAGATGAGAGAGAGAGAGATGAGAGATGAGAGAGAGAGAGAG 540
Db 481 TGTGAGAGTACCAAGAGATGAGAGAGAGAGAGAGATGAGAGATGAGAGAGAGAGAG 540
Qy 541 AAGCCCAATATCATCCAGACAGGAGACCGAGTACACCCGATCCACCTTAGCTGA 597
Db 541 AAGCCCAATATCATCCAGACAGGAGACCGAGTACACCCGATCCACCTTAGCTGA 597

```

RESULT 2
US-09-782-953-2
Sequence 2, Application US/09782953
Patent No. US20020150953A1

```

GENERAL INFORMATION:
APPLICANT: WILLIAMS, R. SANDERS
APPLICANT: ROTHERMEL, BEVERLY
TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
FILE REFERENCE: UTSID:674P21
CURRENT APPLICATION NUMBER: US/09/782,953
CURRENT FILING DATE: 2001-02-13
PRIOR APPLICATION NUMBER: 60/216,601
PRIOR FILING DATE: 2000-07-07
NUMBER OF SEQ ID NOS: 27
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 597
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(594)
US-09-782-953-2

```

Query Match 86.4%; Score 515.8; DB 10; Length 597;
Best Local Similarity 94.4%; Pred. No. 2,3e-151;
Matches 535; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

```

Qy 31 AGCTCCCTGATGCTGTGTGTCGCAACGATGATGCTTTCAGCGAAAGTGAAGACGGGCC 90
Db 31 AGCTCCCTGATGCTGTGTGTCGCAACGATGATGCTTTCAGCGAAAGTGAAGACGGGCC 90
Qy 91 AAATTTGAATCTCTCTTCAGAAATATGACAAAGACACCACTTTCAGATATTTTAAGAGC 150
Db 91 AAATTTGAATCTCTCTTCAGAAATATGACAAAGACACCACTTTCAGATATTTTAAGAGC 150
Qy 151 TTCAAAAGTGTCCGATTAATCTTGAAGCAACCCCTTATCTGAGCGGATGCGAGCGTGG 210
Db 151 TTCAAAAGTGTCCGATTAATCTTGAAGCAACCCCTTATCTGAGCGGATGCGAGCGTGG 210
Qy 211 CTGACAAAGACCGAGTCTGCGGGAAGAAATGAAGTTGATTTTGGCTCAGACTTTACAC 270
Db 211 CTGACAAAGACCGAGTCTGCGGGAAGAAATGAAGTTGATTTTGGCTCAGACTTTACAC 270

```

```

Qy 271 ATGAGAGTTCAACACTGCTCCGCCCAATCCCGCAAAAGATTCTCATCTCCCTCCG 330
Db 271 ATGAGAGTTCAACACTGCTCCGCCCAATCCCGCAAAAGATTCTCATCTCCCTCCG 330
Qy 331 GCGTCCTCCCGTGGTGGTGAAGCAAGTGAAGATGACACCCCGCTGATAAATTACAT 390
Db 331 GCGTCCTCCCGTGGTGGTGAAGCAAGTGAAGATGACACCCCGCTGATAAATTACAT 390
Qy 391 CTTTATATGCGATCTTCAAGCTGGGGCCAGAGAGAAATGAACTGATGACAGCGACA 450
Db 391 CTTTATATGCGATCTTCAAGCTGGGGCCAGAGAGAAATGAACTGATGACAGCGACA 450
Qy 451 GACCCCACTCCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 510
Db 451 GACCCCACTCCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 510
Qy 511 GAGAGAGATGAGAGAGAAATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 570
Db 511 GAGAGAGATGAGAGAGAAATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 570
Qy 571 GAGTACACACCGATCCACCTTAGCTGA 597
Db 571 GAGTACACACCGATCCACCTTAGCTGA 597

```

RESULT 3
US-09-782-953-1
Sequence 1, Application US/09782953
Patent No. US20020150953A1

```

GENERAL INFORMATION:
APPLICANT: WILLIAMS, R. SANDERS
APPLICANT: ROTHERMEL, BEVERLY
TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
FILE REFERENCE: UTSID:674P21
CURRENT APPLICATION NUMBER: US/09/782,953
CURRENT FILING DATE: 2001-02-13
PRIOR APPLICATION NUMBER: 60/216,601
PRIOR FILING DATE: 2000-07-07
NUMBER OF SEQ ID NOS: 27
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 599
TYPE: DNA
ORGANISM: Mus musculus
US-09-782-953-1

```

Query Match 84.5%; Score 504.4; DB 10; Length 599;
Best Local Similarity 99.6%; Pred. No. 8,4e-148;
Matches 516; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```

Qy 4 GATTTTGAAGACTTTAGTCAATTTTATGCTTCTGATGCTTGTGTGTGTGTGTGTGTGT 63
Db 4 GATTTTGAAGACTTTAGTCAATTTTATGCTTCTGATGCTTGTGTGTGTGTGTGTGTGT 63
Qy 64 GCTTCAGGAAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 123
Db 64 GCTTCAGGAAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 123
Qy 124 GACACCACTTCCAGTATTTTAAGAGCTTCAACGCTGTCCGATTAATCTTACAGCAACCC 183
Db 124 GACACCACTTCCAGTATTTTAAGAGCTTCAACGCTGTCCGATTAATCTTACAGCAACCC 183
Qy 184 TTAATGCGAGCGGATGCGAGCTGCGGTGACAAAGAGAGAGAGAGAGAGAGAGAGAG 243
Db 184 TTAATGCGAGCGGATGCGAGCTGCGGTGACAAAGAGAGAGAGAGAGAGAGAGAGAG 243
Qy 244 AAGTTATTTTGTCTCAGACTTTACATAGAGAGTTCAACCTGCTCCGCCCAATCC 303
Db 244 AAGTTATTTTGTCTCAGACTTTACATAGAGAGTTCAACCTGCTCCGCCCAATCC 303
Qy 304 GACAAAGTTCTCATCTCCCTCCGCTCTCTCTCCGTTGGCTGAAAGAAATAGAA 363
Db 304 GACAAAGTTCTCATCTCCCTCCGCTCTCTCTCCGTTGGCTGAAAGAAATAGAA 363

```

Db 382 GACAAACAGTTCTCTCCCTCCGCTCTCTCTCCGTTGGTGAACAAGTAGAA 441
Qy 364 GATGCCACCCCGTCATATAATAGATCTTTTATATGCAATCTCCAGCTGGGGCCAGGA 423
Db 442 GATGCCACCCCGTCATATAATAGATCTTTTATATGCAATCTCCAGCTGGGGCCAGGA 501
Qy 424 GAGAAGTATGAATGTCATGACGACGACAGCCCACTCCCACTGGTGGTCCAGCTGTGT 483
Db 502 GAGAAGTATGAATGTCATGACGACGACAGCCCACTCCCACTGGTGGTCCAGCTGTGT 561
Qy 484 GAGAGTACCAAGAGAAATGAGGAGGAAGAGAGAGAT 521
Db 562 GAGAGTGACCAAGAGAAATGAGGAGGAAGAGAGAT 599

RESULT 4
US-09-782-953-11
; Sequence 11, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: CALCIENEURIN INTERACTING PROTEIN (MCIP)
; CURRENT APPLICATION NUMBER: US/09/782,953
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 2331
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (144)..(734)
US-09-782-953-11

Query Match 78.8%; Score 470.6; DB 10; Length 2331;
Best Local Similarity 87.9%; Pred. No. 6.6e-137;
Matches 525; Conservative 0; Mismatches 69; Indels 3; Gaps 1;
Qy 1 ATGCAATTTAGGACCTTACGTTACAAATTTAGCTCCCTGATTCCTGTGTGGCAACAGT 60
Db 144 ATGCAATTTAGGAACTTTAACTACAGTTTATAGCTCCCTGATTCCTGTGTGGCAACAGT 203
Qy 61 GATGTCTTTCAGGAAAGTGAGACGAGGCGCAAAATTTGAATCCCTCTTCAGACATATGAC 120
Db 204 GATATCTTCAGCGAAGTGAAACGAGGCGCAAAATTTGATGCTCTTTAGGAGGTATGAC 263
Qy 121 AAGGACACACCTTCCAGTATTTAAGAGCTTCAAAAGTGTCCGGATAACTTTCAGCAAC 180
Db 264 AAGGACATCACCTTTCAGTATTTAAGAGCTTCAAAAGAGTCAAGATTAATTTTCAGCAAC 323
Qy 181 CCCTATTCGACCCCGTCAGCTGCGGCTGCAAGACGAGTTCCTGGGGAAGGAA 240
Db 324 CCCTTCTCGCAGCAGATGCCAGGCTCCAGCTGCAATAGACTGAGTTTCTGGGAAGGAA 383
Qy 241 ATGAAGTTGTATTTGCTCAGACTTTTACATAGGAAGTTTCACTCGGCTCCGCCCAT 300
Db 384 ATGAAGTTATATTTGCTCAGACTTTACATAGGAAGTTCACCTCGGCTCCGCCCAAT 443
Qy 301 CCGGACAAACAGTTCTCACTCCCTCCGCTCTCTCCGCTTGGCTGGGAAACAAGTA 360
Db 444 CCAGACAAGCAGTTCTGTATCTCCCTCCGCTCTCCGCTGAGTGGATGGAAACAAGT 503
Qy 361 GAAGATGCCACCCCGTCATATAATAGGATCTTTATATGCAATCTCCAGCTGGGGCA 420
Db 504 GAAGATGCCACCCCGTCATATAATAGTATCTTTATATGCAATCTCCAGCTGGGGCA 563
Qy 421 GGAGAGAAGTATGAATGATGCGGACAGACAGACCCCACTCCCACTGGTGTGTCCACGTG 480

Db 564 GGGGAAAAGTATGAATTTGCAAGCAGGAGCTGACACCACTCCAGCGTGGTCCATGTA 623
Qy 481 TGTGAGAGTACCAACAGAGATGAGGAGGAGAGAGATGAGAGATGAAGAGACCC 540
Db 624 TGTGAGAGTATCAAGAGAAAGGAGGAGGAA---ATGGAAGAATGAGGAGACCT 680
Qy 541 AAGCCCAAAATCATCCAGACACGAGACCGGAGTACACACCGATCCACCTTAGCTGA 597
Db 681 AAGCCCAAAATCATCCAGACGAGGAGGCGGAGTACACCGGATCCACCTCAGCTGA 737
RESULT 5
US-10-044-090-255
; Sequence 255, Application US/10044090
; Patent No. US20020137081A1
; GENERAL INFORMATION:
; APPLICANT: Olga Bandman
; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
; FILE REFERENCE: PA-0028 US
; CURRENT APPLICATION NUMBER: US/10/044,090
; CURRENT FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: PERL Program
; SEQ ID NO 255
; LENGTH: 2355
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020137081A1 5511889CB1
US-10-044-090-255

Query Match 78.8%; Score 470.6; DB 12; Length 2355;
Best Local Similarity 87.9%; Pred. No. 6.6e-137;
Matches 525; Conservative 0; Mismatches 69; Indels 3; Gaps 1;
Qy 1 ATGCAATTTAGGACCTTACGTTACAAATTTAGCTCCCTGATTCCTGTGTGGCAACAGT 60
Db 151 ATGCAATTTAGGAACTTTAACTACAGTTTATAGCTCCCTGATTCCTGTGTGGCAACAGT 210
Qy 61 GATGTCTTTCAGGAAAGTGAGACGAGGCGCAAAATTTGAATCCCTCTTCAGACATATGAC 120
Db 211 GATATCTTTCAGGAAAGTGAACACGAGGCGCAAAATTTGAGTCCCTCTTTAGGAGCTATGAC 270
Qy 121 AAGGACACACCTTCCAGTATTTAAGAGCTTCAAAAGTGTCCGGATAACTTTCAGCAAC 180
Db 271 AAGGACATCACCTTTCAGTATTTAAGAGCTTCAAAAGAGCTTCAAAAGAGTCAAGATTAACCTTCAGCAAC 330
Qy 181 CCCTATCTGACGCGAGTCCAGGCTGCGGCTGCAACAGACCGAGTTCTTGGGAAGGAA 240
Db 331 CCCTTCTCGCAGCAGATGCCAGGCTCCAGCTGCAATAGACTGAGTTTCTGGGAAGGAA 390
Qy 241 ATGAAGTTGTATTTTGTCTCAGACTTTTACATAGGAAGTTTCACTCGGCTCCGCCCAAT 300
Db 391 ATGAAGTTATATTTTGTCTCAGACCTTTACATAGGAAGCTCACACCTGGCTCCGCCCAAT 450
Qy 301 CCGGACAAACAGTTCTTCATCTCCCTCCGCTCTCTCCGCTTGGCTGGAACAAGTA 360
Db 451 CCAGACAAGCAGTTTCTGATCTCCCTCCGCTCTCCGCTGAGTGGATGGAAACAAGT 510
Qy 361 GAAGATGCCACCCCGTCATATAATTAAGATCTTTTATATGCAATCTCCCAAGCTGGGGCA 420
Db 511 GAAGATGCCACCCCGTCATATAATTAAGATCTTTTATATGCAATCTCCCAAGCTGGGGCA 570
Qy 421 GGAGAGAAGTATGAATGATGCGGACAGACAGACCCCACTCCCACTGGTGTGTCCACGTG 480
Db 571 GGGGAAAAGTATGAATTGCACGAGCAGCTGACACCACTCCCAAGCTGGTGTGTCCATGTA 630
Qy 481 TGTGAGAGTACCAAGAGAAATGAGGAGGAGAGAGATGAGAGAGATGAAGAGACCC 540
Db 631 TGTGAGAGTATCAAGAGAAAGGAGGAGGAGGAA---ATGGAAGAATGAGGAGACCT 687


```
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 347
; LENGTH: 2358
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-925-302-347

Query Match      66.2%; Score 395.4; DB 10; Length 2358;
Best Local Similarity 80.9%; Pred. No. 2.1e-113;
Matches 483; Conservative 2; Mismatches 108; Indels 4; Gaps 2;

QY 1 ATGGAATTTAGGCACTTTAGCTACAAATTTAGCTCCCTGATTCGTTGTGGCAACCAT 60
   |||||
Db 114 ATGAGAGAGTGGACCTGCGAGGACCTCCCGAGCGCCACCATCGCTGTCACTGGACCCG 173
   |||||

QY 61 GATGTTCTTCAGCGAAAGTGAGACCGAGCGCCAAATTTGAATCCCTCTTCAGAAACATATGAC 120
   |||||
Db 174 CGGTGTTCTGGACGCGCTGTGCGGGGCCAAATTTGAGTCCCTCTTTAGACGTATGAC 233
   |||||

QY 121 AAGGACACCACTTCAGTATTTTAAAGCTTCAAGCTGTCGCGATPAAATCTCAGCAAC 180
   |||||
Db 234 AAGGACATCACTTTCAAGTATTTTAAAGCTTCAAGCTGTCGCGATPAAATCTCAGCAAC 293
   |||||

QY 181 CCCTTATCTCAGCGGATGCCAGCTGCGGCTGCACAGACCGAGTTCCTGGGGAAGGAA 240
   |||||
Db 294 CCCTTCTCCGAGCAGATGCCAGCTCCAGCTGCATAAGACTGAGTTTCTGGGAAAGGAA 353
   |||||

QY 241 ATGAAGTTGATTTTGTCTCAGACTTTACACATAGGAAGTTTACACCTGGCTCCGCCCAAT 300
   |||||
Db 354 ATGAAGTTATATTTGCTCAGACCTTTACACATAGGAAGTTTACACCTGGCTCCG-CANAT 412
   |||||

QY 301 CCGGACAAACAGTTCCTATCTCCCTCCGCGCTCTCTCTCCGTTGGCTGGAAACAAGTA 360
   |||||
Db 413 CCAGACAAAGCAGTTTCTGATCTCCCTCCGCGCTCTCTCCGSCAGTGGGATGGAACAAGTG 472
   |||||

QY 361 GAAGATCCACCCCGCTCATAAATTTAGATCTTTTATATGCCATCTCCAGCTGGGGCCA 420
   |||||
Db 473 GAAGATCGGACCCAGCTCATAAATTTAGATCTTTTATATGCCATCTCCAGCTGGGGCCA 532
   |||||

QY 421 GGAGAGAAGTATGAATCTGCAATGAGCGACAGACCCCACTCCAGTGTGGTGGTCCACGTG 480
   |||||
Db 533 GGGGAAAAGTATGAATTTGACGAGCGAGCTGACACCACTCCAGCTGGTGGTCCATGTA 592
   |||||

QY 481 TGTGAGAGTACCAAGAGATGAGGAGGAGAGAGAGATGAGAGAGATGAAGAGACCC 540
   |||||
Db 593 TGTGAGAGTATCAAGAGAGAGGAGGAGAGAGAGAGAA---ATGGAAGAAGATGAGAGACCT 649
   |||||

QY 541 AAGCCCAAAATCATCCAGACACCGAGACCGGAGTATACACCGATCCACCTTAGCTGA 597
   |||||
Db 650 AAGCCCAAAATTTATCCAGACAGAGGCGCGAGTACACCGGATCCACCTCAGCTGA 706
   |||||

RESULT 9
US-09-954-456-497
; Sequence 497, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc
; FILE REFERENCE: 689290-76
```

```
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 497
; LENGTH: 3184
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-954-456-497

Query Match      35.9%; Score 214.2; DB 10; Length 3184;
Best Local Similarity 66.1%; Pred. No. 9.8e-57;
Matches 331; Conservative 0; Mismatches 158; Indels 12; Gaps 1;

QY 21 CTACATTTTACTCCCTGATTCCTGTTGTGGCAACGATGATGTTCTTCAGCGAAAGTGA 80
   |||||
Db 210 CTGTGATGTTTCCACTCTGCTGTTGCTGTGTGGTGTGTCGAGGTCTTTTACCAATCAGGA 269
   |||||

QY 81 GACCAGGGGCAAAATTTGAATCCCTCTTCAGAAACATATGACAAAGGACACCACTTCCAGTA 140
   |||||
Db 270 GGTTAAGGAAAATTTGGGGGACTGTTTCGGACTTATGATGACTGTGTGACGTTCCAGCT 329
   |||||

QY 141 TTTTAAGAGCTTCAAACGTGTCCGGATAAACTTCAGCAACCCCTTTATCTGCAGCCGATGC 200
   |||||
Db 330 ATTTAAGAGTTTTCAGACGTTGTCGTATAACTTCAGCAATCTTAAATCTGCAGCCGAGC 389
   |||||

QY 201 CAGGTCGCGCTGCACAGACCGAGTTCCTGGGGAAGGAAATGAAGTTGTATTTTGTCTCA 260
   |||||
Db 390 TAGGATAGAGCTTCATGAACCCCAATTCAGAGGGGAAAAATTTAAAGCTCTACTTTGCACA 449
   |||||

QY 261 GACT-----TTACACATAGGAAGTTTCACACCTGGCTCCGCCCAATCCCGACAA 308
   |||||
Db 450 GGTTCAGACTCCAGAGACAGATGGAGCAAACTGCATCTGGCTCCACCCAGCTTCCCAA 509
   |||||

QY 309 ACAGTTTCTCATCTCCCTCCGCGCTCTCTCTCCCTTGGCTGGAAACAAAGTAGAAGATGC 368
   |||||
Db 510 ACAGTTTCTCATCTCCGCGCTCTCTCTCCCTTGGCTGGAAACAAAGTAGAAGATGC 569
   |||||

QY 369 CACCCCGCTCATAAATTTAGATCTTTTATATGCCATCTCCAAGCTGGGCGCCAGGAGAA 428
   |||||
Db 570 CACGCGAGTCTCAACTATGACTCTCTCTATGCTGTGGCCAAACTAGGACCCAGGAGAA 629
   |||||

QY 429 GTATGAATGTCATGCGAGCAGACAGCCCACTCCAGTGTGGTGGTCCAGCTGTGTGAGAG 488
   |||||
Db 630 GTATGAGCTCCATGCGAGGAGTGTGTCCACCCCAAGTGTCTGTCGTCAGCTGTGCGACAG 589
   |||||

QY 489 TGACCAAGAGATGAGGAGGA 509
   |||||
Db 690 TGACATAGAGGAAGAGAGGA 710
   |||||

RESULT 10
US-09-782-953-17
```

Query Match	35.9%	Score 214.2;	DB 10;	Length 3184;
Best Local Similarity	66.1%	Pred. No. 9.8e-57;		
Matches 331; Conservative	0;	Mismatches 158;	Indels 12;	Gaps 1;

```

FILE REFERENCE: UTSD:674P21
CURRENT APPLICATION NUMBER: US/09/782,953
CURRENT FILING DATE: 2001-02-13
PRIOR APPLICATION NUMBER: 60/216,601
PRIOR FILING DATE: 2000-07-07
NUMBER OF SEQ ID NOS: 27
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 8
LENGTH: 594
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(591)
US-09-782-953-8

Query Match      33.7%; Score 201.4; DB 10; Length 594;
Best Local Similarity 62.8%; Pred. No. 3,7e-53;
Matches 360; Conservative 0; Mismatches 186; Indels 27; Gaps 2;

QY 21 CTACAAATTTTAGCTCCCGATGTCGTGTGGGCAAAAGCATATGCTTCACAGGAATGA 80
D 21 CTGGAATGTTTCCACTCGTGCTGCTGTGTGTGGATGTGGAGGCTTTTACCAATCAGGA 80
QY 81 GACCAAGGCGCAAAATTTTGATCCCTCTTCAGAACATATGACAAAGCAACCACTTTCAGTA 140
D 81 GGTTAAGAAAATTTTCAGAGGACGTGTTCCGAGCCATGATGATATGTGAGCTTCCAGCT 140
QY 141 TTTTAAAGCTTCAAAAGCTGTCCGGATTAACCTTCAGCAACCCCTTATCTGCAGCCGATGC 200
D 141 GTTTAAGATTTTCGAGCGGGTTCCAAATTAATTCAGCCATCCCAATCTGCAGCCGATGC 200
QY 201 CAGGCTGGGCTGCACAAAGACCGAGTTCTCTGGGGAAGAAATGAACTGTATTTTGCTCA 260
D 201 CCGGATAGAGCTTCATGAGACTCAGTTCTAGAGGGAAGAACTTAATACTTACTTGCCCA 260
QY 261 G-----ACTTACACATAGGAAGTTTCAACCTGGCTCCGCCCAATCCCGCA 308
D 261 GGTCCAGACCCCAAGACAGACAGATGAGAGAACAACTGATTTGGCACCTCCACAGCTGCCA 320
QY 309 ACAGTTCTCATCTTCCCTCCGCGCTCTCTCCCTGCGTTGGCTGGAACAAAGTAAAGATGC 368
D 309 ACAAGTCTCATCTTCACTCCCTTCTATCTCTCTGTTGGCTGGAAGCTTATCAGCGATGC 380
QY 369 CACCCCGCTCATTAATTTCGATCTTTTATATGCCATCTTCCAAAGCTGGGCGCAGAGAGA 428
D 369 CACACCACTCTCTCAACTTGAACCTTCTTATGCTGTGCGCAAACTGAGACCCAGGAGAGA 440
QY 429 GATGAACTGCATGATGAGGAGACAGACCCCACTCCCACTGTTGGTGGTCCAGTGTGAGAG 488
D 429 ATATAGCTGCACGCTTGAACTGATCTTACACCAAGCGCTGCTGTGCTCATGTGTGACAG 500
QY 489 TGACCAAGAGATGAGAGGAAGAAAGAGATGAGAGAAATGAAGAGACCCCAAGCCCA 548
D 501 CGACATGAGAGGAGAGGA-----GGACCCAAAGACTTCCCCCAAGCCAA 545
QY 549 AATCATCCAGACACGGAACCGGAAGTACACACC 581
D 546 AATCATTCAGACCCCGCGCTCCGCGGCTCCACCC 578

RESULT 12
US-09-782-953-23
; Sequence 23, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674P21
; CURRENT APPLICATION NUMBER: US/09/782,953
; CURRENT FILING DATE: 2001-02-13

```


; PRIOR APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 720
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(637)
US-09-782-953-23

Query Match 32.0%; Score 191.2; DB 10; Length 720;
Best Local Similarity 63.2%; Pred. No. 6.4e-50;
Matches 320; Conservative 0; Mismatches 168; Indels 18; Gaps 1;
QY 4 GATTTAGGACTTTAGCTACAATTTTGTCTCCCTGATTCCTTGTGTGGCAACGATGAT 63
Db 56 GAGATGATGATTTAGTGATCTGCTACCTCACTTTTGTCTTGTGAGCGTCCATGAAGCA 115
QY 64 GTCTTACGCAAGTGTAGACACCGGCAAAATTTGAATCCCTCTTCAGAACATATGACAAG 123
Db 116 GTGTTTGGGCAACGAGACGAGAGGAAGATTTGAAGCACTCTTCACCATCTATGATGAC 175
QY 124 GACACCACTTCCAGTATTTTGAAGCTTCAAGAGTGTCCGGATTAACCTTCAGCAACCCC 183
Db 176 CAGTTACTTTTTCAGCTGTTTAAAGCTTTTGAAGAGTCAAGATAAATTTTCAGCAACCT 235
QY 184 TTATCTGACGCGATGCGGCTGCACAGACCGAGTTCTGTGGGAGGAAGATG 243
Db 236 GAAGCGGACAGAGCGCAATAGAACTCCAGAAACAGACTTCAATGGGCAAGACTA 295
QY 244 AAGTTGATTTTGTCTAGACTTTTACACATAGGAAGTTTCACACCTGGCTCCGCCCAATCCC 303
Db 296 AAGCTATATTTTGCACAG-----TCTATCTCTCGCGCCCGCCGCT 337
QY 304 GACAAACAGTTCTCATCTCCCTCCGCTCTCTCCGCTCTCTCCGCTGGTGGGCAAGTAGAA 363
Db 338 GTCAAGCAGTTCTCATCTCCCTCCGCTCTCTCCGCTGGTGGGCAAGTAGAA 397
QY 364 GATGCCACCCCGTCAATAATTTAGCATCTTTTATATGCAATCTCCAGCTGGGCGCCAGGA 423
Db 398 GATGCGAGCTCTTATATATATGATTTACTCTGTCTGTTTCCAAATTTGGGACCCAGGA 457
QY 424 GAGAAGTATGAATGTGATGACGACGACACCCCACTCCCACTGGTGGTGGTCCAGCTGTGT 483
Db 458 GAGAAATATGAATCTTCAAGCGGCAACAGAGTCCACACCCCGCTGGTGGTCTATGCTGT 517
QY 484 GAGAGTCAACCAAGATGAGGAGGA 509
Db 518 GAAAGTGAACCTGAAGAGGAAGA 543

RESULT 13

US-09-782-953-20
; Sequence 20, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)
; FILE REFERENCE: US/09/782,953
; CURRENT APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 828
; TYPE: DNA

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (23)..(745)
US-09-782-953-20

Query Match 32.0%; Score 191.2; DB 10; Length 828;
Best Local Similarity 62.4%; Pred. No. 6.9e-50;
Matches 323; Conservative 0; Mismatches 183; Indels 12; Gaps 1;
QY 4 GATTTAGGACTTTAGCTACAATTTTGTCTCCCTGATTCCTTGTGTGGCAACGATGAT 63
Db 134 GAGATGATGATTTAGTGATCTGCTACCTCACTTTTGTCTTGTGAGCGTCCATGAAGCA 193
QY 64 GTCTTACGCAAGTGTAGACACCGGCAAAATTTGAATCCCTCTTCAGAACATATGACAAG 123
Db 194 GTGTTTGGGCAACGAGACGAGAGGAAGATTTGAAGCACTCTTCACCATCTATGATGAC 253
QY 124 GACACCACTTCCAGTATTTTGAAGCTTCAAACTGTCCGGATTAACCTTCAGCAACCCC 183
Db 254 CAGTTACTTTTTCAGCTGTTTAAAGCTTTTGAAGAGTCAAGATAAATTTTCAGCAACCT 313
QY 184 TTATCTGACGCGATGCGGCTGCACAGACCGAGTTCTTCAGCAACGAGTTCCTGGGAGGAAGATG 243
Db 314 GAAGCGGCAACGAGCGCAATAGAACTCCACGAAACAGACTTCAATGGGCAAGACTA 373
QY 244 AAGTTGATTTTGTCTCAG-----ACTTTACACATAGGAAGTTTCACACCTGGCT 291
Db 374 AAGCTATATTTTGCACAGTGTCCGCGAGGAGTCCGCGCAAGTCTCTATCTCTG 433
QY 292 CGGCCAATTCGCAACAAACAGTTCTCATCTCCCTCCGCTCTCTCTCCCTGGCTGG 351
Db 434 CGGCCCAACCTGTCAAGCAGTCTCTCATCTCCCTCCGCTCTCTCCCTGGCTGG 493
QY 352 AAACAAGTAGAGATGCCACCCCGCTCATATAATTTAGCATCTTTTATATGCAATCTCCCAAG 411
Db 494 AAGCAGAGCAAGATGCGATGCTCTTATAATTTATGATTTACTCTGTCTGTTTCCAA 553
QY 412 CTGGGCGCAGGAGAGGAAGTATGAATGCAATGCAAGCAGACACCCCACTCCCACTGGTG 471
Db 554 TTGGGACCGAGAGGAATATGAATTTCAAGCGGCAACAGAGTCCACCCAGCGTGGTG 613
QY 472 GTCCACGTTGTGAGAGTGACCAAGAGAAATGAGGAGGA 509
Db 614 GTTCATGCTGTGAAGTGAAGTCAAGAGGAAGA 651

RESULT 14

US-09-864-761-10388/c
; Sequence 10388, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Acomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30

```

PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO: 10388
LENGTH: 412
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AP000054.1
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.6
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.96
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.67
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.62
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.67
US-09-864-761-10388
```

```

Query Match      24.2%; Score 144.6; DB 10; Length 412;
Best Local Similarity 89.1%; Pred. No. 1.7e-35;
Matches 156; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 87 GGGCAATTGATCCCTCTTCAGACATATGACAGACACACCTTCAGATTATTTAA 146
DB 411 GGGCAATTGATCCCTCTTCAGACATATGACAGACATACCTTCAGATTATTTAA 352
QY 147 GAGCTTCAACGTCGCGATTAACCTTATTCAGCCGATGCCAGGCT 206
DB 351 GAGCTTCAACGTCGCGATTAACCTTATTCAGCCGATGCCAGGCT 292
QY 207 GCGGCTGCACAGACCGAGTTCTGGGGAAGAAATGAAGTTGATTTGCTCAG 261
DB 291 CCAGCTGCATTAAGCTGATTTCTGGGGAAGAAATGAAGTTGATTTGCTCAG 237

RESULT 15
US-09-864-761-2064/c
Sequence 2064, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aomic-X-1
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US/09/864,761
PRIOR APPLICATION NUMBER: US 60/180,312
```

```

PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO: 2064
LENGTH: 446
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AP000122.1
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.1
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 4.1
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.7
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.3
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.2
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.4
US-09-864-761-2064
```

```

Query Match      24.2%; Score 144.6; DB 10; Length 446;
Best Local Similarity 89.1%; Pred. No. 1.7e-35;
Matches 156; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 87 GGGCAATTGATCCCTCTTCAGACATATGACAGACACCACTTCAGATTATTTAA 146
DB 445 GGGCAATTGATCCCTCTTCAGACATATGACAGACATACCTTCAGATTATTTAA 386
QY 147 GAGCTTCAACGTCGCGATTAACCTTATTCAGCCGATGCCAGGCT 206
DB 385 GAGCTTCAACGTCGCGATTAACCTTATTCAGCCGATGCCAGGCT 326
QY 207 GCGGCTGCACAGACCGAGTTCTGGGGAAGAAATGAAGTTGATTTGCTCAG 261
DB 325 CCAGCTGCATTAAGCTGATTTCTGGGGAAGAAATGAAGTTGATTTGCTCAG 271

Search completed: December 14, 2002, 20:01:42
```

Sun Dec 15 08:38:26 2002

Job time: 24.3119 secs

us-09-782-953-5.rnpb

Page 9

THIS PAGE BLANK (usf70)


```

QY 41 GUCYVA1THrPheGlnLeuPheLysSerPheArgValArgIleAsnPheSerHis 60
   : : : : :
Db 310 GACTGTGAGCTTCCAGCTATTAAAGTTTCACAGCGTCCGATATAAATTGAGCAAT 369
QY 61 PROLYSERRA1A1AARG1A1ARG1LEGLUeuh1SG1UTHrGlnPheArgGlyLys 80
   CCTAATCTGACGCCGAGCTAGATAGAGCTTCATGAACCCCAATTGAGAGGAAAAAA 429
Db 370
QY 81 LEULYSLEUTHrPheAlaGlnValGlnThrProGlnUTHrAspGlyAspLysLeuHis 100
   LEULYSLEUTHrPheAlaGlnValGlnThrProGlnUTHrAspGlyAspLysLeuHis 100
Db 430 TTAAGCTCTACTTTCACAGGTTCAAGCTCCAGAGACAGATGAGACAAACTGCACCTTG 489
QY 101 ALAPROG1INPRO1ALALYSG1INPheLeu1LSErPRO1SER1SER1PRO1VAL1GLY 120
   490 GCTCACCCCTCAGCTCCGAAACAGTTTCATCTCGCCCCCTCTCCCTCCACCTGTGGC 549
Db 490
QY 121 TRPLYSER1LSErAspAlaThrProValLeuAsnTYrAspLeuLeuTYrAlaValAla 140
   550 TGGCAGCCCATCAAGCATGCCAGCGCAGCTCCCACTATGAGCTCTCTATGCTGTGGC 609
Db 550
QY 141 LYSLEUG1YPRO1GLY1ULYrTYrGluLeuHisAlaGlyTHrGluSerThrProSerVal 160
   610 AAACCTAGACACAGAGAGAAATGATGAGCTTCATGACGAGACTGAGTCCACCCCAAGTGC 669
Db 610
QY 161 VALVALHISVAL1CYBAspSerAspMetGluGluGluAspProLYrSThrSerProLYs 180
   670 GTCTGTCACGTGTGCGACAGTGCATAGAGAGAAAGAACCCAAAGACTTCCCAAG 729
Db 670
QY 181 PROLYS1LE1LEGLNTHrArgArg 188
   730 CCAAAATCATCCAGCGGTGATAGA 753
Db 730

```

RESULT 2

```

US-09-724-676A-15329
; Sequence 15329, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 15329
; LENGTH: 2411
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676A-15329

```

Alignment Scores:

```

Pred. No.: 6,39e-87 Length: 2411
Score: 950.00 Matches: 180
Percent Similarity: 98.40% Conservative: 5
Best Local Similarity: 95.74% Mismatches: 3
Query Match: 91.26% Indels: 0
DB: 5 Gaps: 0

```

US-09-782-953-9 (1-197) x US-09-724-676A-15329 (1-2411)

```

QY 1 METPRO1ALPROSER1METAspCYBAspVal1SerThrLeuValAlaCYBVal1AspVal 20
   190 ATGCCACCCCTAGCATGAGTGTGATGTTCCACTCTGTGTGCTGTGTGTGTGTGAGTGC 249
Db 190
QY 21 GLUVALPheThrAsnGlnGluValLysGluLysPheGluGluLysPheArgThrTYrAsp 40
   250 GAGGTCTTTACCAATCAGAGGTTAAAGAAAATTTGGGGACGTGTTCCGACTATGAT 309
Db 250
QY 41 GUCYVA1THrPheGlnLeuPheLysSerPheArgValArgIleAsnPheSerHis 60
   : : : : :
Db 310 GACTGTGAGCTTCCAGCTATTAAAGTTTCACAGCGTCCGATATAAATTGAGCAAT 369
QY 61 PROLYSERRA1A1AARG1A1ARG1LEGLUeuh1SG1UTHrGlnPheArgGlyLys 80
   : : : : :

```

```

Db 370 CCTAATCTGACGCCGAGCTAGATAGAGCTTCATGAACCCCAATTGAGAGGAAAAAA 429
QY 81 LEULYSLEUTHrPheAlaGlnValGlnThrProGlnUTHrAspGlyAspLysLeuHis 100
   430 TTAAGCTCTACTTTCACAGGTTTCACAGCTCCAGAGACAGATGAGACAAACTGCACCTTG 489
Db 430
QY 101 ALAPROG1INPRO1ALALYSG1INPheLeu1LSErPRO1SER1SER1PRO1VAL1GLY 120
   490 GCTCACCCCTCAGCTCCGAAACAGTTTCATCTCGCCCCCTCTCCCTCCACCTGTGGC 549
Db 490
QY 121 TRPLYSER1LSErAspAlaThrProValLeuAsnTYrAspLeuLeuTYrAlaValAla 140
   550 TGGCAGCCCATCAAGCATGCCAGCGCAGCTCCCACTATGAGCTCTCTATGCTGTGGC 609
Db 550
QY 141 LYSLEUG1YPRO1GLY1ULYrTYrGluLeuHisAlaGlyTHrGluSerThrProSerVal 160
   610 AAACCTAGACACAGAGAGAAATGATGAGCTTCATGACGAGACTGAGTCCACCCCAAGTGC 669
Db 610
QY 161 VALVALHISVAL1CYBAspSerAspMetGluGluGluAspProLYrSThrSerProLYs 180
   670 GTCTGTCACGTGTGCGACAGTGCATAGAGAGAAAGAACCCAAAGACTTCCCAAG 729
Db 670
QY 181 PROLYS1LE1LEGLNTHrArgArg 188
   730 CCAAAATCATCCAGCGGTGATAGA 753
Db 730

```

RESULT 3

```

US-10-290-438-3/c
; Sequence 3, Application US/10290438
; GENERAL INFORMATION:
; APPLICANT: Loring, Jeanne F.
; APPLICANT: Tingley, Debora W.
; APPLICANT: Edwards, Carla W.
; APPLICANT: Streeter, David G.
; TITLE OF INVENTION: DOWN SYNDROME CRITICAL REGION 1-LIKE PROTEIN
; FILE REFERENCE: PC-0013-1CIP
; CURRENT APPLICATION NUMBER: US/10/290,438
; FILING DATE: 2002-11-06
; PRIOR APPLICATION NUMBER: 09/614,474
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PERL Program
; SEQ ID NO: 3
; LENGTH: 1021
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 219435_Rn.1
US-10-290-438-3

```

Alignment Scores:

```

Pred. No.: 2,28e-83 Length: 1021
Score: 910.00 Matches: 192
Percent Similarity: 94.63% Conservative: 2
Best Local Similarity: 93.66% Mismatches: 3
Query Match: 87.42% Indels: 8
DB: 6 Gaps: 0

```

US-09-782-953-9 (1-197) x US-10-290-438-3 (1-1021)

```

QY 1 METPRO1ALPROSER1METAspCYBAspVal1SerThrLeuValAlaCYBVal1AspVal 20
   855 ATGCCACCCCTAGCATGAGTGTGATGTTCCACTCTGTGTGCTGTGTGTGTGTGAGTGC 796
Db 855
QY 21 GLUVALPheThrAsnGlnGluValLysGluLysPheGluGluLysPheArgThrTYrAsp 40
   795 GAGGTCTTTACCAATCAGAGGTTAAAGAAAATTTGAAGGACGTGTTCCGACTATGAT 736
Db 795
QY 41 GUCYVA1THrPheGlnLeuPheLysSerPheArgValArgIleAsnPheSerHis 60
   : : : : :
Db 735 GATGTGACGTTTCAGCTGTTAAAGTTTCGACGCGGTTCGGAATTAATTGAGCCA 676

```

Qy 60 sProLysSerAlaAlaArg-AlaArgIle-GluLeuHisGlu-ThrGlnPheArg-GlyL 79
 Db 675 CCCCCAAGCTGCGAGCCGGTCCCGATAGGAGCTTCATGTGGACCCAGTTTCATGAGGA 616
 Qy 79 ySLysLeuLysLeuTyrPheAlaGlnValGlnThrProGluThrAspGly-AspLysLeu 98
 Db 615 AGAAGCTGAAACTTACCTTTTCACAGGTCCAGACCCCGACAGACAGATGGTAGACAAACTG 556
 Qy 99 His-LeuAlaProGlnProAlaLysGlnPheLeuLysSerProProSerSerProPr 118
 Db 555 CAGTTTGGCACCCCGACAGCTGCGCAACAGTTCCTCATCTCACCCCTTCATCTCTCC 496
 Qy 118 oValGlyTrrLysProLysSerAspAlaThrProValLeuAsn-TyrAspLeuLeuTyrA 138
 Db 495 CTTTGGCTGGGAAGCTATCAGCGATGCCACACCGATGCTCAACGTACGACCTCTTTATG 436
 Qy 138 laValAlaLysLeuGlyProGlyGluLysTyrGluLeuHisAlaGlyThrGluSerThrP 158
 Db 435 CCGTGGCCAACTAGGACAGAGAGAAATATGAGCTGCATGCGGGAACAGTGTACAC 376
 Qy 158 roSerValValHisValCysAspSerAspMetGluGluGluAspProLysThrS 178
 Db 375 CGAGCGTTGTCGTGACCGTGTGACAGCGACTTGGAGGAGGAGGATCCAAAGACTT 316
 Qy 178 erProLysProLysIleleGlnThrArgArgProGlyLeuProProSerValSerAsn 197
 Db 315 CCCCCAAGCCAAATATCATCCAGACCCCGCGTCTCGCCTGCCACCCCTCGGTGTCCAAC 257

RESULT 4

US-10-290-438-1
 ; Sequence 1, Application US/10290438
 ; GENERAL INFORMATION:
 ; APPLICANT: Loring, Jeanne F.
 ; APPLICANT: Tingley, Debora W.
 ; APPLICANT: Edwards, Carla M.
 ; APPLICANT: Streeter, David G.
 ; TITLE OF INVENTION: DOWN SYNDROME CRITICAL REGION 1-LIKE PROTEIN
 ; CURRENT APPLICATION NUMBER: US/10/290,438
 ; PRIOR FILING DATE: 2002-11-06
 ; PRIOR APPLICATION NUMBER: 09/614,474
 ; PRIOR FILING DATE: 2000-07-11
 ; NUMBER OF SEQ ID NOS: 11
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 1
 ; LENGTH: 3159
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; OTHER INFORMATION: Incyte ID No: 247500.5
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: 216, 3091, 3093, 3103
 ; OTHER INFORMATION: a, t, c, g, or other

US-10-290-438-1

Alignment Scores:
 Pred. No.: 3,26e-82 Length: 3159
 Score: 905.00 Matches: 171
 Percent Similarity: 95.72% Conservative: 8
 Best Local Similarity: 91.44% Mismatches: 8
 Query Match: 86.94% Indels: 0
 DB: 6 Gaps: 0

US-09-782-953-9 (1-197) x US-10-290-438-1 (1-3159)

Qy 11 SerThrLeuValAlaCysValValAspValGluValPheThrAsnGlnGluValLysGlu 30
 Db 389 AACTCGTTGTTTCGTCGATGTCACCGATGATGTTTGAAGGAGAGAGAGAGAGAA 448
 Qy 31 LysPheGluGlyLeuPheArgThrTyrAspGluCysValThrPheGlnLeuPheLysSer 50

Db 449 AAATTTGAGGAGCTGTTTCGGACTTATGACTGTGTGACGTTCCAGCTATTATTAAAGAGT 508
 Qy 51 PheArgValArgIleAsnPheSerHisProLysSerAlaAlaArgAlaArgIleGlu 70
 Db 509 TTCCAGACGTGTCCGTATAAACTTCAGCAATCTCTAAATCTGCAGCCGAGCTAGGATAGAG 568
 Qy 71 LeuHisGluThrGlnPheArgGlyLysLeuLysLeuTyrPheAlaGlnValGlnThr 90
 Db 569 CTTTCATGAACCCCAATTCAGAGGGAAAAAATAAAGCTCTACTTTGCACAGGTTTCAGACT 628
 Qy 91 ProGluThrAspGlyAspLysLeuHisLeuAlaProProGlnProAlaLysGlnPheLeu 110
 Db 629 CCAGACAGATGGAGACAACTGCACCTTGGCTCCACCCCGAGCTGCCAAACAGTTTCTC 688
 Qy 111 IleSerProProSerSerProProValGlyTrrLysProLysSerAspAlaThrProVal 130
 Db 689 ATCTGCCCCCTTCTCTCCACCTGTGTGCGCAGCCCATCAACGATGCCACGCGCAGTC 748
 Qy 131 LeuAsnTyrAspLeuLeuTyrAlaValAlaLysLeuGlyProGlyGluLysTyrGluLeu 150
 Db 749 CTCACACTATGACTCTCTATGCTGTGGCCAACTAGGACCGAGAGAGATATGAGCTC 808
 Qy 151 HisAlaGlyThrGluSerThrProSerValValHisValCysAspSerAspMetGlu 170
 Db 809 CATGCGGAGCTGAGTCCACCCCAAGTGTGCTGCGTGCAGCTGTGCACAGTCACATAGAG 868
 Qy 171 GluGluGluAspProLysThrSerProLysProLysIleleGlnThrArgArgProGly 190
 Db 869 GAAGAAGAGAGAGCCCAAGACTTCCCAAGCCAAAAATCATCCAAACTCGGCGCTCTGCG 928
 Qy 191 LeuProProSerValSerAsn 197
 Db 929 CTGCCACCCCTCGTGTCCAAAC 949

RESULT 5

US-09-724-676-12748
 ; Sequence 12748, Application US/09724676
 ; GENERAL INFORMATION:
 ; APPLICANT: Compugen LTD
 ; TITLE OF INVENTION: Variants of alternative splicing
 ; FILE REFERENCE: 129181.4 Compugen
 ; CURRENT APPLICATION NUMBER: US/09/724,676
 ; CURRENT FILING DATE: 2000-11-28
 ; NUMBER OF SEQ ID NOS: 97222
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 12748
 ; LENGTH: 1931
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (11)-(11)
 ; OTHER INFORMATION: n is a,c,g, or t
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (18)-(18)
 ; OTHER INFORMATION: n is a,c,g, or t
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (23)-(23)
 ; OTHER INFORMATION: n is a,c,g, or t

Alignment Scores:

Pred. No.: 8.64e-56 Length: 1931
 Score: 641.00 Matches: 120
 Percent Similarity: 79.68% Conservative: 29
 Best Local Similarity: 64.17% Mismatches: 30
 Query Match: 61.58% Indels: 8
 DB: 5 Gaps: 2

US-09-782-953-9 (1-197) x US-09-724-676-12748 (1-1931)

QY 11 SerThrLeuValAlaCysValAlaSerValAlaGluValPheThrAsnGlnGluValLeuGlu 30
 DB 93 AGCTCCCTGATTCCTCTGTCGCAACAGTGATATCTTCAGCGAAAGGAAACGAGGCC 152
 QY 31 LysPheGlnGluLeuPheArgThrTyraSPGluCysValThrPheGlnLeuPheLysSer 50
 DB 153 AAATTGAGTCCCTCTTATGACGATATGACAAAGACATCACCTTTCAGATTTTAAAGGC 212
 QY 51 PheArgValAlaArgIleAsnPheSerHisProLysSerAlaAlaArgIleGlu 70
 DB 213 TTCAACGAGTCAGATTAATTAATTCAGCAACCCCTTCTCCGACGAGATCCAGGCTCCAG 272
 QY 71 LeuHisGlnThrGlnPheArgGlyLysLeuLysLeuLysLeuTyraPheAlaGlnValGlnThr 90
 DB 273 CTGCATTAAGACTGAGTTCTGCGAAAGAAATGATTAATTTTGTCTCAGACCTTAAC 332
 QY 91 ProGluThrAspGlyAspLysLeuHisLeuAlaProProGlnProAlaLysGlnPheLeu 110
 DB 333 ATAGCAAGCTCA-----CACCTGGCTCCGCCAAATCCAGACAAAGAGTTTCTG 380
 QY 111 ILeSerProProSerSerProProValGlyTyrLysProIleSerAspAlaThrProVal 130
 DB 381 ATCTCCCTCCGCTCTCCGACGATGGAATGAAACAAAGTGGAAGATCCGACCCAGTC 440
 QY 131 LeuAsnTyrAspLeuLeuTyraAlaValAlaLysLeuGlyProGlyGluLysTyrGluLeu 150
 DB 441 ATAAACTATGATCTCTTATATGCACTCCAGCTGGGGCCAGGGGAAAGATATGATTTG 500
 QY 151 HisAlaGlyThrGlnSerThrProSerValValHisValCysAspSerAspMetGlu 170
 DB 501 CACGCGACGATGACACACACTCCACGCTGGTGTCCATGATATGAGATGATCAAGAG 560
 QY 171 GlnGlnGluAspProLysThrSer-----ProLysProLysIleIleGlnThr 186
 DB 561 AAGGAGAAAGAAAGAAATGAAAGATGAGAGACCTAAGCCAAATTTATCCAGACC 620
 QY 187 ArgArgProGlyLeuProPro 193
 DB 621 AGGAGCCGAGTACACGCCG 641

RESULT 6

US-09-724-676A-12748
 ; Sequence 12748, Application US/09724676A
 ; GENERAL INFORMATION:
 ; APPLICANT: Compugen LTD
 ; TITLE OF INVENTION: Variants of alternative splicing
 ; FILE REFERENCE: 129181.4 Compugen
 ; CURRENT APPLICATION NUMBER: US/09/724,676A
 ; NUMBER OF SEQ ID NOS: 97222
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 12748
 ; LENGTH: 1931
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (11)..(11)
 ; OTHER INFORMATION: n is a,c,g, or t
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (18)..(18)
 ; OTHER INFORMATION: n is a,c,g, or t
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (23)..(23)
 ; OTHER INFORMATION: n is a,c,g, or t
 ; US-09-724-676A-12748

Alignment Scores:

Pred. No.: 8 64e-56 Length: 1931
 Score: 641.00 Matches: 120
 Percent Similarity: 79.68% Conservative: 29

Best Local Similarity: 64.17% Mismatches: 30
 Query Match: 61.58% Indels: 8
 DB: 5 Gaps: 2

US-09-782-953-9 (1-197) x US-09-724-676A-12748 (1-1931)

QY 11 SerThrLeuValAlaCysValAlaSerValAlaGluValPheThrAsnGlnGluValLeuGlu 30
 DB 93 AGCTCCCTGATTCCTCTGTCGCAACAGTGATATCTTCAGCGAAAGGAAACGAGGCC 152
 QY 31 LysPheGlnGluLeuPheArgThrTyraSPGluCysValThrPheGlnLeuPheLysSer 50
 DB 153 AAATTGAGTCCCTCTTATGACGATATGACAAAGACATCACCTTTCAGATTTTAAAGGC 212
 QY 51 PheArgValAlaArgIleAsnPheSerHisProLysSerAlaAlaArgIleGlu 70
 DB 213 TTCAACGAGTCAGATTAATTAATTCAGCAACCCCTTCTCCGACGAGATCCAGGCTCCAG 272
 QY 71 LeuHisGlnThrGlnPheArgGlyLysLeuLysLeuLysLeuTyraPheAlaGlnValGlnThr 90
 DB 273 CTGCATTAAGACTGAGTTCTGCGAAAGAAATGATTAATTTTGTCTCAGACCTTAAC 332
 QY 91 ProGluThrAspGlyAspLysLeuHisLeuAlaProProGlnProAlaLysGlnPheLeu 110
 DB 333 ATAGCAAGCTCA-----CACCTGGCTCCGCCAAATCCAGACAAAGAGTTTCTG 380
 QY 111 ILeSerProProSerSerProProValGlyTyrLysProIleSerAspAlaThrProVal 130
 DB 381 ATCTCCCTCCGCTCTCCGACGATGGAATGAAACAAAGTGGAAGATCCGACCCAGTC 440
 QY 131 LeuAsnTyrAspLeuLeuTyraAlaValAlaLysLeuGlyProGlyGluLysTyrGluLeu 150
 DB 441 ATAAACTATGATCTCTTATATGCACTCCAGCTGGGGCCAGGGGAAAGATATGATTTG 500
 QY 151 HisAlaGlyThrGlnSerThrProSerValValHisValCysAspSerAspMetGlu 170
 DB 501 CACGCGACGATGACACACACTCCACGCTGGTGTCCATGATATGAGATGATCAAGAG 560
 QY 171 GlnGlnGluAspProLysThrSer-----ProLysProLysIleIleGlnThr 186
 DB 561 AAGGAGAAAGAAAGAAATGAAAGATGAGAGACCTAAGCCAAATTTATCCAGACC 620
 QY 187 ArgArgProGlyLeuProPro 193
 DB 621 AGGAGCCGAGTACACGCCG 641

RESULT 7

US-09-724-676-12756
 ; Sequence 12756, Application US/09724676
 ; GENERAL INFORMATION:
 ; APPLICANT: Compugen LTD
 ; TITLE OF INVENTION: Variants of alternative splicing
 ; FILE REFERENCE: 129181.4 Compugen
 ; CURRENT APPLICATION NUMBER: US/09/724,676
 ; NUMBER OF SEQ ID NOS: 97222
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 12756
 ; LENGTH: 2050
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (7)..(7)
 ; OTHER INFORMATION: n is a,c,g, or t
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (47)..(47)
 ; OTHER INFORMATION: n is a,c,g, or t
 ; US-09-724-676-12756

Alignment Scores:

Pred. No.: 9 35e-56 Length: 2050

Score: 641.00 Matches: 120
 Percent Similarity: 79.68% Conservativity: 29
 Best Local Similarity: 64.17% Mismatches: 30
 Query Match: 61.58% Indels: 8
 DB: 5 Gaps: 2

US-09-782-953-9 (1-197) x US-09-724-676-12756 (1-2050)

```

Qy 11 SerThrLeuValAlaCysValValAspValGluValPheThrAsnGlnGluValLysGlu 30
Db 212 AGCTCCCTGATGCTGTGTGCAACAGTGTATCTTCAGCGAAGTGAACACAGGGCC 271
Qy 31 LysPheGluGlyLeuPheArgThrTyAspGluCysValThrPheGlnLeuPheLysSer 50
Db 272 AAATTTGAGTCCCTCTTTAGGACGTATGACAAAGACATCACCTTTTCAGTATTTTAAGAGC 331
Qy 51 PheArgValAlaArgIleAsnPheSerHisProLysSerAlaAlaAArgAlaArgIleGlu 70
Db 332 TTCAACAGAGTCAGATTAACCTTCAGCAACCCCTTCCTCCGACGAGATGCCAGGCTCCAG 391
Qy 71 LeuHisGluThrGlnPheArgGlyLysLeuLysLeuTyThrPheAlaGlnValGlnThr 90
Db 392 CTGCATAAGACTGAGTTCTCGGAAAGCAATGAAGTTATATTTTGTCTCAGACCTTACAC 451
Qy 91 ProGluThrAspGlyAspLysLeuHisLeuAlaProProGlnProAlaLysGlnPheLeu 110
Db 452 ATAGAGAGCTCA-----CACCTGGCTCCGCCAAATCCAGACAAAGCAGTTTCTG 499
Qy 111 IleSerProProSerSerProProValGlyTyrLysProLysProLysSerAlaThrProVal 130
Db 500 ATCTCCCTCCCGCTCTCCGCAAGTGGATGGAAACAGTGAAGATGGACCCAGTC 559
Qy 131 LeuAsnTyrAspLeuLeuTyrAlaValAlaLysLeuGlyProGlyGluLysTyrGluLeu 150
Db 560 ATAAACTATGATCTCTTATATGCCATCTCCAAGCTGGGGCCAGGGGAAAGTATGAATTG 619
Qy 151 HisAlaGlyThrGluSerThrProSerValValHisValCysAspSerAspMetGlu 170
Db 620 CACGACGAGTACACCATCTCCAGCGTGGTGTCCATGTATGTGAGAGTATCAAGAG 679
Qy 171 GluGluAspProLysThrSer-----ProLysProLysIleIleGlnThr 186
Db 680 AAGAGGAAGAGGAAATGGAAGAAATGAGGAGACCTTAAGCCAAANATTATCCAGACC 739
Qy 187 ArgArgProGlyLeuProPro 193
Db 740 AGGAGCGCGAGTACACGCG 760

```

RESULT 8

```

US-09-724-676-12756
; Sequence 12756, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 12756
; LENGTH: 2050
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (7)..(7)
; OTHER INFORMATION: n is a,c,g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (47)..(47)
; OTHER INFORMATION: n is a,c,g, or t
US-09-724-676A-12756

```

Alignment Scores:

Pred. No.: 9,358-56 Length: 2050
 Score: 641.00 Matches: 120
 Percent Similarity: 79.68% Conservativity: 29
 Best Local Similarity: 64.17% Mismatches: 30
 Query Match: 61.58% Indels: 8
 DB: 5 Gaps: 2

US-09-782-953-9 (1-197) x US-09-724-676A-12756 (1-2050)

```

Qy 11 SerThrLeuValAlaCysValValAspValGluValPheThrAsnGlnGluValLysGlu 30
Db 212 AGCTCCCTGATGCTGTGTGCAACAGTGTATCTTCAGCGAAGTGAACACAGGGCC 271
Qy 31 LysPheGluGlyLeuPheArgThrTyAspGluCysValThrPheGlnLeuPheLysSer 50
Db 272 AAATTTGAGTCCCTCTTTAGGACGTATGACAAAGACATCACCTTTTCAGTATTTTAAGAGC 331
Qy 51 PheArgValAlaArgIleAsnPheSerHisProLysSerAlaAlaAArgAlaArgIleGlu 70
Db 332 TTCAACAGAGTCAGATTAACCTTCAGCAACCCCTTCCTCCGACGAGATGCCAGGCTCCAG 391
Qy 71 LeuHisGluThrGlnPheArgGlyLysLeuLysLeuTyThrPheAlaGlnValGlnThr 90
Db 392 CTGCATAAGACTGAGTTCTCGGAAAGCAATGAAGTTATATTTTGTCTCAGACCTTACAC 451
Qy 91 ProGluThrAspGlyAspLysLeuHisLeuAlaProProGlnProAlaLysGlnPheLeu 110
Db 452 ATAGAGAGCTCA-----CACCTGGCTCCGCCAAATCCAGACAAAGCAGTTTCTG 499
Qy 111 IleSerProProSerSerProProValGlyTyrLysProLysProLysSerAlaThrProVal 130
Db 500 ATCTCCCTCCCGCTCTCCGCAAGTGGATGGAAACAGTGAAGATGGACCCAGTC 559
Qy 131 LeuAsnTyrAspLeuLeuTyrAlaValAlaLysLeuGlyProGlyGluLysTyrGluLeu 150
Db 560 ATAAACTATGATCTCTTATATGCCATCTCCAAGCTGGGGCCAGGGGAAAGTATGAATTG 619
Qy 151 HisAlaGlyThrGluSerThrProSerValValHisValCysAspSerAspMetGlu 170
Db 620 CACGACGAGTACACCATCTCCAGCGTGGTGTCCATGTATGTGAGAGTATCAAGAG 679
Qy 171 GluGluAspProLysThrSer-----ProLysProLysIleIleGlnThr 186
Db 680 AAGAGGAAGAGGAAATGGAAGAAATGAGGAGACCTTAAGCCAAANATTATCCAGACC 739
Qy 187 ArgArgProGlyLeuProPro 193
Db 740 AGGAGCGCGAGTACACGCG 760

```

RESULT 9

```

US-09-724-676-12747
; Sequence 12747, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 12747
; LENGTH: 2297
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (11)..(11)
; OTHER INFORMATION: n is a,c,g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (18)..(18)
; OTHER INFORMATION: n is a,c,g, or t

```

FEATURE:
 NAME/KEY: misc feature
 LOCATION: (23)-(23)
 OTHER INFORMATION: n is a,c,g, or t
 US-09-724-676-12747

Alignment Scores:

Score:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
641.00	2297	120	29	30	8	2
Percent Similarity: 79.68%						
Best Local Similarity: 64.17%						
Query Match: 61.58%						

US-09-782-953-9 (1-197) x US-09-724-676-12747 (1-2297)

```

Oy 11 SerThrLeuValAlaCyValAlaAspValAlaPheThrAsnGlnGluValIleGlu 30
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Db 93 AGCTCCCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
Oy 31 LysPheGluGlyLeuPheArgThrTyraAspGluCysValThrPheGlnLeuPheLysSer 50
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Db 153 AAATTTGAGTCCCTCTTATGAGAGCTTATGACAGACATCACTTTCAGTATTTAAGAC 212
Oy 51 PheArgValAlaArgIleAsnPheSerHisProLysSerAlaAlaArgIleGlu 70
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Db 213 TTCAACGAGTCAGATTAACCTTCAGCAACCCCTTCTCCGAGCAGATGCCAGGCTCCAG 272
Oy 71 LeuHisGluThrGlnPheArgGlyIleValLeuLysLeuTyrrPheAlaGlnValGlnThr 90
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Db 273 CTGCATTAAGCTGATGTTCTGCGAAGAAATGAAAGTTATATTGCTCAAGCTTACAC 332
Oy 91 ProGluThrAspGlyAspLysLeuHisLeuAlaProProGlnProAlaLysGlnPheLeu 110
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Db 333 ATAGGAAGCTCA-----CACCTGGCTCCGCCAAATCCACAGACAGCTTCTG 380
Oy 111 IleSerProProSerSerProProValGlyTyrrLysProLysSerAspAlaThrProVal 130
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Db 381 ATCTCCCTCCCGCTCTCCGCGAGTGGATGGAAACAAATGAGATGCAACCCAGCTC 440
Oy 131 LeuAsnTyrrAspLeuLeuTyrrAlaValAlaLysLeuGlyProGlyGluTyrrGlnLeu 150
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Db 441 ATAACTATGATCTCTTATATGCCATCTCCAACTGGGGCAGGGAGAAAGTATGAAATTG 500
Oy 151 HisAlaGlyThrGluSerThrProSerValValHisValCysAspSerAspMetGlu 170
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Db 501 CACGCGACGACTGACACACACTCCCAAGGCTGCTCATGTATGTAGAGTATCAAGAG 560
Oy 171 GluGluGluAspProLysThrSer-----ProLysProLysIleIleGlnThr 186
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Db 561 AAGGAGGAAGAAAGAAATGGAAGAAATGAGAGACTTAAGCCAAAATTTATCCAGACC 620
Oy 187 ArgArgProGlyLeuProPro 193
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Db 621 AGGAGCGCGAGTACACGCCG 641
  
```

RESULT 10

US-09-724-676A-12747
 Sequence 12747, Application US/09724676A
 GENERAL INFORMATION:
 APPLICANT: Compugen LTD
 TITLE OF INVENTION: Variants of alternative splicing
 FILE REFERENCE: 129181.4 Compugen
 CURRENT APPLICATION NUMBER: US/09/724,676A
 CURRENT FILING DATE: 2000-11-28
 NUMBER OF SEQ ID NOS: 97222
 SOFTWARE: PatentIn version 3.2
 SEQ ID NO 12747
 LENGTH: 2297
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: misc feature

LOCATION: (11)-(11)
 OTHER INFORMATION: n is a,c,g, or t
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (18)-(18)
 OTHER INFORMATION: n is a,c,g, or t
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (23)-(23)
 OTHER INFORMATION: n is a,c,g, or t
 US-09-724-676A-12747

Alignment Scores:

Score:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
641.00	2297	120	29	30	8	2
Percent Similarity: 79.68%						
Best Local Similarity: 64.17%						
Query Match: 61.58%						

US-09-782-953-9 (1-197) x US-09-724-676A-12747 (1-2297)

```

Oy 11 SerThrLeuValAlaCyValAlaAspValAlaPheThrAsnGlnGluValIleGlu 30
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Db 93 AGCTCCCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
Oy 31 LysPheGluGlyLeuPheArgThrTyraAspGluCysValThrPheGlnLeuPheLysSer 50
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Db 153 AAATTTGAGTCCCTCTTATGAGAGCTTATGACAGACATCACTTTCAGTATTTAAGAC 212
Oy 71 LeuHisGluThrGlnPheArgGlyIleValLeuLysLeuTyrrPheAlaGlnValGlnThr 90
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Db 273 CTGCATTAAGCTGATGTTCTGCGAAGAAATGAAAGTTATATTGCTCAAGCTTACAC 332
Oy 91 ProGluThrAspGlyAspLysLeuHisLeuAlaProProGlnProAlaLysGlnPheLeu 110
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Db 333 ATAGGAAGCTCA-----CACCTGGCTCCGCCAAATCCACAGACAGCTTCTG 380
Oy 111 IleSerProProSerSerProProValGlyTyrrLysProLysSerAspAlaThrProVal 130
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Db 381 ATCTCCCTCCCGCTCTCCGCGAGTGGATGGAAACAAATGAGATGCAACCCAGCTC 440
Oy 131 LeuAsnTyrrAspLeuLeuTyrrAlaValAlaLysLeuGlyProGlyGluTyrrGlnLeu 150
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Db 441 ATAACTATGATCTCTTATATGCCATCTCCAACTGGGGCAGGGAGAAAGTATGAAATTG 500
Oy 151 HisAlaGlyThrGluSerThrProSerValValHisValCysAspSerAspMetGlu 170
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Db 501 CACGCGACGACTGACACACACTCCCAAGGCTGCTCATGTATGTAGAGTATCAAGAG 560
Oy 171 GluGluGluAspProLysThrSer-----ProLysProLysIleIleGlnThr 186
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Db 561 AAGGAGGAAGAAAGAAATGGAAGAAATGAGAGACTTAAGCCAAAATTTATCCAGACC 620
Oy 187 ArgArgProGlyLeuProPro 193
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Db 621 AGGAGCGCGAGTACACGCCG 641
  
```

RESULT 11

US-10-240-965-60
 Sequence 60, Application US/10240965
 GENERAL INFORMATION:
 APPLICANT: INCYTE GENOMICS, INC.
 APPLICANT: SHIFMAN, DOV
 APPLICANT: SOMOGYI, ROLAND
 APPLICANT: LAWN, RICHARD M.
 APPLICANT: SEILHAMER, JEFFREY J.
 APPLICANT: PORTER, Gordon J.
 APPLICANT: MIXITA, Thomas

; APPLICANT: TAI, Julie
 ; TITLE OF INVENTION: GENES EXPRESSED IN FOAM CELL DIFFERENTIATION
 ; FILE REFERENCE: PA-0025 PCT
 ; CURRENT APPLICATION NUMBER: US/10/240,965
 ; CURRENT FILING DATE: 2002-10-04
 ; PRIOR APPLICATION NUMBER: 60/195,106
 ; PRIOR FILING DATE: 2000-04-05
 ; NUMBER OF SEQ ID NOS: 276
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 60
 ; LENGTH: 2348
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; OTHER INFORMATION: Incyte ID No: 042176.5
 US-10-240-965-60

Alignment Scores:
 Pred. No.: 2,31e-54 Length: 2348
 Score: 628.00 Matches: 120
 Percent Similarity: 79.26% Conservative: 29
 Best Local Similarity: 63.83% Mismatches: 30
 Query Match: 60.33% Indels: 9
 DB: 6 Gaps: 2

US-09-782-953-9 (1-197) x US-10-240-965-60 (1-2348)

Qy 11 SerThrLeuValAlaCysValValAspValGluValPheThrAsnGlnGluValLysGlu 30
 Db 181 AGCTCCCTGATTGCTGTGTGGCAACAGTATATCTTACGGAAAGTGAACAGGGCC 240
 Qy 31 LysPheGluGlyLeuPheArgThrTyAspGluCysValThrPheGlnLeuPheLysSer 50
 Db 241 AAATTGAGTCCCTCTTTAGACGCTATGACAAGGACATCACCTTTTCAGTATTTTAAGAGC 300
 Qy 51 PheArgValArgGileAenPheSerHisProLysSerAlaAlaArgAlaArgGileGlu 70
 Db 301 TTCAACGAGTCAGATTAACCTTCAGCAACCCCTTCTCCGAGGAGATGCCAGGCTCCAG 360
 Qy 71 LeuHisGluThrGlnPheArgGlyLysLeuLysLeuTyPheAlaGlnValGlnThr 90
 Db 361 CTGCATAAGACTGAGTTCTTGGGAAGAAATGAAGTTATATTTGCTCAGACTTACAC 420
 Qy 91 ProGluThrAspGlyAspLysLeuHisLeu-AlaProProGlnProAlaLysGlnPheLe 110
 Db 421 ATAGAAAGCTCA-----CACCTGGGCTCCGCCAAATCCAGACAAGCAGTTTCT 468
 Qy 110 uileSerProProSerSerProProValGlyTyrPheProLysAspAlaThrProVa 130
 Db 469 GATCTCCCTCCCGCTCTCCGCGAGTGGATGGAACAAGTGAAGATGCCAGCCAGT 528
 Qy 130 LeuAsnTyAspLeuLeuTyAlaValAlaLysLeuGlyProGlyLysTyGluLe 150
 Db 529 CATAAATATGATCTCTTATATGCCATCTCCAGCTGGGCCAGGGGAAAGATGATGAT 588
 Qy 150 uHisAlaGlyThrGluSerThrProSerValValHisValCysAspSerAspMetG 170
 Db 589 GCACGCGCCAGTACACCACTCCAGCGTGTGTGTCATGTATGAGAGTATGATCAAGA 648
 Qy 170 uGluGluGluAspProLysThrSer-----ProLysProLysIleIleGlnTh 186
 Db 649 GAAGGAGGAGAGAGAGAAATGAAGAATGAGGAGACCTTAAGCCAAATATATCCAGAC 708
 Qy 186 rArgArgProGlyLeuProPro 193
 Db 709 CAGGAGCGCGAGTACAGCGC 730

RESULT 12

US-09-724-676-12741
 ; Sequence 12741, Application US/09724676
 ; GENERAL INFORMATION:
 ; APPLICANT: Compugen LTD

; TITLE OF INVENTION: Variants of alternative splicing
 ; FILE REFERENCE: 129181.4 Compugen
 ; CURRENT APPLICATION NUMBER: US/09/724,676
 ; CURRENT FILING DATE: 2000-11-28
 ; NUMBER OF SEQ ID NOS: 97222
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 12741
 ; LENGTH: 1893
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-724-676-12741

Alignment Scores:

Pred. No.: 2,47e-54 Length: 1893
 Score: 626.50 Matches: 121
 Percent Similarity: 74.63% Conservative: 29
 Best Local Similarity: 60.20% Mismatches: 34
 Query Match: 60.18% Indels: 17
 DB: 5 Gaps: 3

US-09-782-953-9 (1-197) x US-09-724-676-12741 (1-1893)

Qy 6 MetAspCysAspVal-----SerThrLeuValAlaCys 16
 Db 13 ATTGACTGCGAGATGAGGAGGTGGACCTCGAGGACCTGCCCGCCGCCACCATCGCTGT 72
 Qy 17 ValValAspValGluValPheThrAsnGlnGluValLysGluLysPheGluGlyLeuPhe 36
 Db 73 CACCTGGACCCCGCGCTGTCTGTGACGCGCTGTCCCGGGCCAAATTTGAGTCCCTCTTT 132
 Qy 37 ArgThrTyAspGluCysValThrPheGlnLeuPheLysSerPheArgArgValArgile 56
 Db 133 AGGAGCTATGACAAGGACATCACCTTTTCAGTATTTTAAGAGCTTCAACGAGTCAAGATA 192
 Qy 57 AsnPheSerHisProLysSerAlaAlaArgAlaArgGileGluLeuHisGluThrGlnPhe 76
 Db 193 AACCTTCAGCAACCCCTTCTCCGAGCAGATGCCAGGCTCCAGCTGCATAAGACTGAGTTT 252
 Qy 77 ArgGlyLysLeuLysLeuTyPheAlaGlnValGlnThrProGluThrAspGlyAsp 96
 Db 253 CTGGGAAGGAATGAAGTTATATTTGCTCAGACTTACACATAGGAAGCTCA----- 306
 Qy 97 LysLeuHisLeuAlaProProGlnProAlaLysGlnPheLeuLysSerProProSerSer 116
 Db 307 -----CACCTGGCTCCGCCAAATCCAGACAAGCAGTTTCTGATCTCCCTCCCGCTCT 360
 Qy 117 ProProValGlyTyrPheProLysSerAspAlaThrProValLeuAsnTyAspLeuLeu 136
 Db 361 CGCCAGTGGGATGGAACAAGTGAAGATGCCAGCCAGTCATAAACTATGATCTCTTA 420
 Qy 137 TyrAlaValAlaLysLeuGlyProGlyLysTyGluLeuHisAlaGlyThrGluSer 156
 Db 421 TATGCCATCTCAAGCTGGGCCAGGGGAAAGTATGAATTCAGCAGCCAGCTGACACC 480
 Qy 157 ThrProSerValValHisValCysAspSerAspMetGluGluGluAspProLys 176
 Db 481 ACTCCAGCGTGGTCCATGTATGTGAGAGTGTATCAAGAGAGAGGAGAGAGAGAA 540
 Qy 177 ThrSer-----ProLysProLysIleIleGlnThrArgArgProGlyLeuPro 192
 Db 541 ATGGAAGAATGAGGAGACCTTAAGCCAAATATATCCAGACCCAGGAGCGGAGTACACG 600
 Qy 193 Pro 193
 Db 601 CCG 603

RESULT 13

US-09-724-676A-12741
 ; Sequence 12741, Application US/09724676A
 ; GENERAL INFORMATION:
 ; APPLICANT: Compugen LTD
 ; TITLE OF INVENTION: Variants of alternative splicing
 ; FILE REFERENCE: 129181.4 Compugen

```

; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12741
; LENGTH: 1893
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676A-12741

Alignment Scores:
Pred. No.: 2,476-54      Length: 1893
Score: 626.50           Matches: 21
Percent Similarity: 74.63%  Conservative: 29
Best Local Similarity: 60.20%  Mismatches: 34
Query Match: 60.18%      Indels: 17
DB: 5                   Gaps: 3

US-09-782-953-9 (1-197) x US-09-724-676A-12741 (1-1893)
QY 6 MetAapCCyAaApVal-----SeThrLeuValAlaCys 16
DB 13 ATTGACTGCGAATGAGAGAGTGACCTCGAGAGACCTCCAGCCGACCATCGCTGT 72
QY 17 ValValaApValGluValPheThraEngInGluValLysGluLysPheGluLysLeuPhe 36
DB 73 CACCTGAGACCCCGCGCTGTTCGTGACGCGCTGTGCGGCGCAATTTGAGTCCCTCTT 132
QY 37 ArgThrTyraPrgLucyValThrPheGlnLeuPheLysSerPheArgTyraValArgIle 56
DB 133 AGACAGCTATGACACAGACATCACCTTCAGTATTTCAGCTTCAACGAGTCAGAAATA 192
QY 57 AmphSerHisPrgLysSerAlaAlaArgAlaArgIleGluLeuHisGluThrGlnPhe 76
DB 193 AACTTACGAAACCCCTTCTCCGACAGATGCCAGGCTCCAGCTGATAAAGTGAAGTT 252
QY 77 ArgGlyLeuLeuLeuLysLeuTyrrPheAlaGlnValGlnThrPrgLysLeuPrgLysLeu 96
DB 253 CTGGGAAAGGAATGAATGATATTTGCTCAGACCTTACACATAGGAAGCTCA----- 306
QY 97 LysLeuHisLeuAlaPrgPrgLysPrgLysPrgLysPrgLysPrgLysPrgLysPrgLysPrg 116
DB 307 -----CACCTGCTCCGCGCAATCCAGACACAGCTTCTGATCTCCCTCCGCTCT 360
QY 117 PrgPrgValGlyTyrrPrgPrgLysPrgLysPrgLysPrgLysPrgLysPrgLysPrgLysPrg 136
DB 361 CCCCGAGTGGATGAGAAACAAGTGAAGATGCCAGCCCACTCATTAAGTATGATCTCTTA 420
QY 137 TyraValAlaLysLeuGlyPrgGlyGluLysTyrrGluLeuHisAlaGlyThrGlnSer 156
DB 421 TATGCCATCTCCAAAGCTGGGCGGAGGAAAGTATGCAACGAGGAGCATGACACC 480
QY 157 ThrPrgSerValValValHisValCysAapSerAapMetGluGluGluGluGluGluGlu 176
DB 481 ACTCCAGCTGTGTGCTCATGTATGTGAGATGATCAAGAGGAGGAGGAGGAGGAGGAGGAG 540
QY 177 ThSer-----PrgLysPrgLysIleGlnThrArgPrgPrgLysLeuPrg 192
DB 541 ATGGAAGAATGAGAGACCTAAGCCAAAATTTATCCAGCCAGGAGCGGAGGTACAGC 600

RESULT 14
US-09-724-676-12743
; Sequence 12743, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variance of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; CURRENT FILING DATE: 2000-11-28

```

```

; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12743
; LENGTH: 1839
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676-12743

Alignment Scores:
Pred. No.: 1,146-51      Length: 1839
Score: 600.00           Matches: 115
Percent Similarity: 79.77%  Conservative: 23
Best Local Similarity: 66.47%  Mismatches: 27
Query Match: 57.64%      Indels: 8
DB: 5                   Gaps: 2

US-09-782-953-9 (1-197) x US-09-724-676-12743 (1-1839)
QY 25 AengInGluValLysGluLysPheGluGlyLeuPheArgThrTyraPrgLucyValThr 44
DB 43 AACAGAAATGTTGGGCGCAATTTGAGTCCCTTTAGACGTATGACAGACATCACC 102
QY 45 PheGlnLeuPheLysSerPheArgTyraValArgIleAapPheSerHisPrgLysSerAla 64
DB 103 TTTCAGTATTTCAGAGCTTCAACGAGTCAGATTAACCTTCAGCAACCTTCTCCGCA 162
QY 65 AlaArgAlaArgIleGluLeuHisGluThrGlnPheArgGlyLysLysLeuLysLeuTyrr 84
DB 163 GCAAGTCCAGAGCTCCAGCTCATTAAGACTGATTTCTGGGAAAGAAATGAAGTATAT 222
QY 85 PheAlaGlnValGlnThrPrgLysPrgLysPrgLysPrgLysPrgLysPrgLysPrgLysPrg 104
DB 223 TTTGCTCAGACCTTACATAGGAAGCTCA-----CACCTGCTCCGCGCAAT 270
QY 105 PrgAlaLysGlnPheLeuHisSerPrgPrgPrgPrgPrgPrgPrgPrgPrgPrgPrgPrgPrg 124
DB 271 CCAGCAAGCAAGCTTCTGATCTCCCTCCGCTCCGCGCAAGTGGAGAAACAAGTG 330
QY 125 SerAapAlaThrPrgValLeuAapTyrrAapLeuLeuTyrrAlaValAlaLysLeuGlyPrg 144
DB 331 GAAAGTGGCAAGCCCGCATTAAGATGATCTTATATGCAATCTCCAACTCGGCGCA 390
QY 145 GlyLysTyrrGluLeuHisAlaGlyThrGlnSerThrPrgSerValValHisVal 164
DB 391 GGGGAAAGTAAATTCACGACCGCACTCAACACTCCAGCGTGTGTGATGTA 450
QY 165 CysAapSerAapMetGluGluGluLysPrgLysPrgLysPrgLysPrgLysPrgLysPrgLysPrg 180
DB 451 PGTGAGATGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 510
QY 181 PrgLysIleIleGlnThrArgPrgPrgLysLeuPrgPrg 193
DB 511 CCAAAATATATCCAGACGAGGCGGAGTACACGCCG 549

RESULT 15
US-09-724-676A-12743
; Sequence 12743, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variance of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12743
; LENGTH: 1839
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676A-12743

Alignment Scores:
Pred. No.: 1,146-51      Length: 1839

```

Score:	600.00	Matches:	115
Percent Similarity:	79.77%	Conservative:	23
Best Local Similarity:	66.47%	Mismatches:	27
Query Match:	57.64%	Indels:	8
DB:	5	Gaps:	2

US-09-782-953-9 (1-197) x US-09-724-676A-12743 (1-1839)

QY	25	AenGlnGluValLysGluLysPheGluGlyLeuPheArgThrTyrAspGluCysValThr	44
DB	43	AACAGAATGGTGTGGGCCAAATTTGAGTCCCTCTTTAGACGTATGACAGGACATCAC	102
QY	45	PheGlnLeuPheLysSerPheArgValArgLleAenPheSerHisProLysSerAla	64
DB	103	TTTCAGATATTTTAAGAGCTTCAACAGAGTCAGATAAACTTCAGCAACCCCTTCTCCGCA	162
QY	65	AlaArgAlaArgIleGluLeuHisGluThrGlnPheArgGlyLysLysLeuLysLeuTyr	84
DB	163	GCAGATGCCAGGCTCCAGCTGCATAGACTGAGTTTCTGGGAAAGGAATATGAAGTTATAT	222
QY	85	PheAlaGlnValGlnThrProGluThrAspGlyAspLysLeuHisLeuAlaProProGln	104
DB	223	TTTGCTCAGACCTTACACATAGGAAGCTCA-----CACCTGGCTCGGCCAAAT	270
QY	105	ProAlaLysGlnPheLeuLysSerProProSerSerProProValGlyTyrLysProIle	124
DB	271	CCAGACAGCAGATTCTGATCTCCCTCCCGCTCTCCGCCAGTGGGATGGAAACAAGTG	330
QY	125	SerAspAlaThrProValLeuAenTyrAspLeuLeuTyrAlaValAlaLysLeuGlyPro	144
DB	331	GAAGATCGCACCCCGACTCAATAACTATGATCTCTATATGCCATCTCCAAGCTGGGGCCA	390
QY	145	GlyGluLysTyrGluLeuHisAlaGlyThrGluSerThrProSerValValHisVal	164
DB	391	GGGGAAAGATGATGAATTGCACCGACGACTGACACCACTCCAGCGTGGTGGTCCATGTA	450
QY	165	CysAspSerAspMetGluGluGluAspProLysThrSer-----ProLys	180
DB	451	TGTCAGAGTGTCAACAGAGAAGGAGGAGAGAGGAATATGAAGAATGATGAGGAGACCTAAG	510
QY	181	ProLysIleIleGlnThrArgArgProGlyLeuPro	193
DB	511	CCAAAAATTTATCAGACCCAGGAGCGCGGATACACGCGG	549

Search completed: December 15, 2002, 00:04:31
Job time : 58.489 secs

THIS PAGE BLANK (USPTO)

Qy 131 INYDLVAVAKLGPGEKYLHAGTESTPSVVVHVCSDMEEDPKTS---PKPKIIQT 186
 Db 127 INYDLVAISKLGPEKELHNAATDTTSPVVVHVCSEDEKEEBEMRMRPKPKIIQT 186
 Qy 187 TRPGALPP 193
 Db 187 TRPEYTP 193

RESULT 9
 US-09-782-953-24
 ; Sequence 24, Application US/09782953
 ; Patent No. US20020150953A1
 ; GENERAL INFORMATION:
 ; APPLICANT: WILLIAMS, R. SANDERS
 ; APPLICANT: ROTHERMEL, BEVERLY
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
 ; FILE REFERENCE: UTSD:674P21
 ; CURRENT APPLICATION NUMBER: US/09/782,953
 ; PRIOR FILING DATE: 2001-02-13
 ; PRIOR APPLICATION NUMBER: 60/216,601
 ; NUMBER OF SEQ ID NOS: 27
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 24
 ; LENGTH: 212
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-782-953-24

Query Match 60.9%; Score 633.5; DB 10; Length 212;
 Best Local Similarity 67.9%; Pred. No. 8.5e-55;
 Matches 127; Conservative 17; Mismatches 32; Indels 11; Gaps 2;

Qy 11 STLVAQVDEVEFTNOEVEKEFGLFRTYDECVTFLFKSFRVRINFSHPKSAARARIE 70
 Db 28 TSLFACSVHEAVFEARQEKERFEALFTIYDDQVTFOLFKSFRVRINFSKPEAAARARIE 87
 Qy 71 LHETOPRGKRLKYFAOVQTPETDGDKLHLAPQAKOFLISPPSPVGMKQSEDATPV 130
 Db 88 LHETDFNGQKRLKYFAO-----STLLPQPVKQFLISPPASPPVGMKQSEDAAPV 137
 Qy 131 INYDLVAVAKLGPGEKYLHAGTESTPSVVVHVCSDMEEDPKTSPPKPKIIQTRPG 190
 Db 138 INYDLVAISKLGPEKELHNAATDTTSPVVVHVCSEDEKEEBEMRMRPKPKIIQT 196

Qy 191 LPPSVSN 197
 Db 197 PPTAALN 203

RESULT 10
 US-09-782-953-25
 ; Sequence 25, Application US/09782953
 ; Patent No. US20020150953A1
 ; GENERAL INFORMATION:
 ; APPLICANT: WILLIAMS, R. SANDERS
 ; APPLICANT: ROTHERMEL, BEVERLY
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
 ; FILE REFERENCE: UTSD:674P21
 ; CURRENT APPLICATION NUMBER: US/09/782,953
 ; PRIOR FILING DATE: 2001-02-13
 ; PRIOR APPLICATION NUMBER: 60/216,601
 ; NUMBER OF SEQ ID NOS: 27
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 25
 ; LENGTH: 212
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens

US-09-782-953-25

Query Match 60.9%; Score 633.5; DB 10; Length 212;
 Best Local Similarity 67.9%; Pred. No. 8.5e-55;
 Matches 127; Conservative 17; Mismatches 32; Indels 11; Gaps 2;

Qy 11 STLVAQVDEVEFTNOEVEKEFGLFRTYDECVTFLFKSFRVRINFSHPKSAARARIE 70
 Db 28 TSLFACSVHEAVFEARQEKERFEALFTIYDDQVTFOLFKSFRVRINFSKPEAAARARIE 87
 Qy 71 LHETOPRGKRLKYFAOVQTPETDGDKLHLAPQAKOFLISPPSPVGMKQSEDATPV 130
 Db 88 LHETDFNGQKRLKYFAO-----STLLPQPVKQFLISPPASPPVGMKQSEDAAPV 137
 Qy 131 INYDLVAVAKLGPGEKYLHAGTESTPSVVVHVCSDMEEDPKTSPPKPKIIQTRPG 190
 Db 138 INYDLVAISKLGPEKELHNAATDTTSPVVVHVCSEDEKEEBEMRMRPKPKIIQT 196
 Qy 191 LPPSVSN 197
 Db 197 PPTAALN 203

RESULT 11
 US-09-782-953-6
 ; Sequence 6, Application US/09782953
 ; Patent No. US20020150953A1
 ; GENERAL INFORMATION:
 ; APPLICANT: WILLIAMS, R. SANDERS
 ; APPLICANT: ROTHERMEL, BEVERLY
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
 ; FILE REFERENCE: UTSD:674P21
 ; CURRENT APPLICATION NUMBER: US/09/782,953
 ; PRIOR FILING DATE: 2001-02-13
 ; PRIOR APPLICATION NUMBER: 60/216,601
 ; NUMBER OF SEQ ID NOS: 27
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 6
 ; LENGTH: 198
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 US-09-782-953-6

Query Match 60.8%; Score 632.5; DB 10; Length 198;
 Best Local Similarity 64.4%; Pred. No. 9.7e-55;
 Matches 121; Conservative 24; Mismatches 34; Indels 9; Gaps 2;

Qy 11 STLVAQVDEVEFTNOEVEKEFGLFRTYDECVTFLFKSFRVRINFSHPKSAARARIE 70
 Db 11 SLLIACVANDVDFSESETRAFESLFRYYDKDTTFQYFKSFRVRINFSNPLSADARLR 70
 Qy 71 LHETOPRGKRLKYFAOVQTPETDGDKLHLAPQAKOFLISPPSPVGMKQSEDATPV 130
 Db 71 LHETDFNGQKRLKYFAOQLHIGSS---HLAPVNDKQFLISPPASPPVGMKQSEDAAPV 126
 Qy 131 INYDLVAVAKLGPGEKYLHAGTESTPSVVVHVCSDMEEDPKTSPPKPKIIQ 185
 Db 127 INYDLVAISKLGPEKELHNAATDTTSPVVVHVCSEDEKEEBEMRMRPKPKIIQT 186
 Qy 186 TRPGALPP 193
 Db 187 TRPEYTP 194

RESULT 12
 US-09-782-953-7
 ; Sequence 7, Application US/09782953
 ; Patent No. US20020150953A1
 ; GENERAL INFORMATION:
 ; APPLICANT: WILLIAMS, R. SANDERS
 ; APPLICANT: ROTHERMEL, BEVERLY


```
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674PZ1
; CURRENT APPLICATION NUMBER: US/09/782,953
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-782-953-7

Query Match          60.8%; Score 632.5; DB 10; Length 198;
Best Local Similarity 64.4%; Pred. No. 9.7e-55;
Matches 121; Conservative 24; Mismatches 34; Indels 9; Gaps 2;

QY 11 STLAVCVVDEVTNOEVKEFGLRPTVDECVTFQLPKSPRRVRINFSPKSAARARIE 70
DB 11 SSLIACVANDVSESTRAKFESLFTYDKDITTFQYFKSFKRVRINFSPKSAADARLR 70
QY 71 LHETQFRGKKLYFAQVQTPETDGDKLHLAPPQAKOFLISPPSPVGVKQVEDATPV 130
DB 71 LKTEFLGKEMKLYFAQTLHIGSS----HLAPPNDKQFLISPPSPVGVKQVEDATPV 126
QY 131 LNYDLLYAVAKLGPGKGYELHAGTSTPSVVHVHVCSDMEEEEDPKT-----TPKPKIIQ 185
DB 127 INYDLLYAIKLGPGKGYELHAATDTPSPVVHVHVCSDQNEEEEMERMRPKPKIIQ 186
QY 186 TRRGLPP 193
DB 187 TRPEYTP 194

RESULT 13
US-09-782-953-15
; Sequence 15, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674PZ1
; CURRENT APPLICATION NUMBER: US/09/782,953
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-15

Query Match          59.8%; Score 622; DB 10; Length 197;
Best Local Similarity 63.6%; Pred. No. 1e-53;
Matches 119; Conservative 26; Mismatches 34; Indels 8; Gaps 2;

QY 11 STLAVCVVDEVTNOEVKEFGLRPTVDECVTFQLPKSPRRVRINFSPKSAARARIE 70
DB 11 SATIACHLDPRVFDGLCRKAFESLFTYDKDITTFQYFKSFKRVRINFSPKSAADARLR 70
QY 71 LHETQFRGKKLYFAQVQTPETDGDKLHLAPPQAKOFLISPPSPVGVKQVEDATPV 130
DB 71 LKTEFLGKEMKLYFAQTLHIGSS----HLAPPNDKQFLISPPSPVGVKQVEDATPV 126
QY 131 LNYDLLYAVAKLGPGKGYELHAGTSTPSVVHVHVCSDMEEEEDPKT-----PKPKIIQ 186
DB 127 INYDLLYAIKLGPGKGYELHAATDTPSPVVHVHVCSDQNEEEEMERMRPKPKIIQ 186
QY 187 TRRGLPP 193
DB 187 TRPEYTP 194

RESULT 15
US-09-782-953-3
; Sequence 3, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674PZ1
; CURRENT APPLICATION NUMBER: US/09/782,953
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-782-953-3

Query Match          58.9%; Score 613.5; DB 10; Length 198;
Best Local Similarity 63.3%; Pred. No. 6.9e-53;
```

```
QY 187 TRRGLPP 193
DB 187 TRPEYTP 193

RESULT 14
US-09-782-953-16
; Sequence 16, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674PZ1
; CURRENT APPLICATION NUMBER: US/09/782,953
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-16

Query Match          59.8%; Score 622; DB 10; Length 197;
Best Local Similarity 63.6%; Pred. No. 1e-53;
Matches 119; Conservative 26; Mismatches 34; Indels 8; Gaps 2;

QY 11 STLAVCVVDEVTNOEVKEFGLRPTVDECVTFQLPKSPRRVRINFSPKSAARARIE 70
DB 11 SATIACHLDPRVFDGLCRKAFESLFTYDKDITTFQYFKSFKRVRINFSPKSAADARLR 70
QY 71 LHETQFRGKKLYFAQVQTPETDGDKLHLAPPQAKOFLISPPSPVGVKQVEDATPV 130
DB 71 LKTEFLGKEMKLYFAQTLHIGSS----HLAPPNDKQFLISPPSPVGVKQVEDATPV 126
QY 131 LNYDLLYAVAKLGPGKGYELHAGTSTPSVVHVHVCSDMEEEEDPKT-----PKPKIIQ 186
DB 127 INYDLLYAIKLGPGKGYELHAATDTPSPVVHVHVCSDQNEEEEMERMRPKPKIIQ 186
QY 187 TRRGLPP 193
DB 187 TRPEYTP 193

RESULT 15
US-09-782-953-3
; Sequence 3, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674PZ1
; CURRENT APPLICATION NUMBER: US/09/782,953
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-782-953-3

Query Match          58.9%; Score 613.5; DB 10; Length 198;
Best Local Similarity 63.3%; Pred. No. 6.9e-53;
```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 14, 2002, 21:32:16 ; Search time 32.8887 Seconds
(without alignments)
2368.771 Million cell updates/sec

Title: US-09-782-953-9

Perfect score: 1041

Sequence: 1 MPAPSDCDVSTLVACVDV.....SPAPKIIQTRRGLPPSVSN 197

Scoring table:

	BLOSUM62
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 355320 seqs, 197730502 residues

Total number of hits satisfying chosen parameters: 710640

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09782953/runat_11122002_114431_17465/app.query.fasta_1.1173
-DB=Published Applications NA -OFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS-human40.cdi -LIST=45 -LOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09782953@cgn 1 1 38 @runat_11122002_114431_17465
-NCFU=6 -ICPU=3 -NO LXPY -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:

	1:	2:	3:	4:	5:	6:	7:	8:	9:	10:	11:	12:	13:	14:
/cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:														
/cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:														
/cgn2_6/ptodata/2/pubpna/US05_NEW_PUB.seq:														
/cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:														
/cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:														
/cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:														
/cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:														
/cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:														
/cgn2_6/ptodata/2/pubpna/US03_PUBCOMB.seq:														
/cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:														
/cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:														
/cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:														
/cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:														

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1041	100.0	594	10	US-09-782-953-8 Sequence 8, Appl
2	1007	96.7	3184	10	US-09-954-456-497 Sequence 497, App
3	1007	96.7	3184	10	US-09-782-953-17 Sequence 17, Appl
4	658.5	63.3	828	10	US-09-782-953-20 Sequence 20, Appl

ALIGNMENTS

RESULT 1

US-09-782-953-8

; Sequence 8, Application US/09782953

; Patent No. US20020150953A1

; GENERAL INFORMATION:

; APPLICANT: WILLIAMS, R. SANDERS

; APPLICANT: ROTHERMEL, BEVERLY

; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE

; FILE REFERENCE: US/09782,953

; CURRENT APPLICATION NUMBER: US/09782,953

; CURRENT FILING DATE: 2001-02-13

; PRIOR APPLICATION NUMBER: 60/216,601

; PRIOR FILING DATE: 2000-07-07

; NUMBER OF SEQ ID NOS: 27

; SOFTWARE: Patent In Ver. 2.1

; SEQ ID NO 8

; LENGTH: 594

; TYPE: DNA

; ORGANISM: Mus musculus

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(591)

US-09-782-953-8

Alignment Scores:

Pred. No.:

Score:

6.35e-107

Length:

1041.00

Matches:

594

197

Sequence 11, Appl
Sequence 255, App
Sequence 23, Appl
Sequence 5, Appl
Sequence 14, Appl
Sequence 2, Appl
Sequence 340, Ap
Sequence 337, App
Sequence 1, Appl
Sequence 10388, A
Sequence 1084, Ap
Sequence 736, App
Sequence 809, App
Sequence 27019, A
Sequence 17592, A
Sequence 18808, A
Sequence 1474, Ap
Sequence 361, App
Sequence 422, App
Sequence 422, App
Sequence 422, App
Sequence 536, App
Sequence 4005, Ap
Sequence 302, App
Sequence 522, App
Sequence 3371, Ap
Sequence 9, Appl
Sequence 3, Appl
Sequence 27196, A
Sequence 10575, A
Sequence 105, App
Sequence 1740, Ap
Sequence 13, Appl
Sequence 5722, Ap
Sequence 1, Appl
Sequence 1691, Ap
Sequence 1691, Ap
Sequence 583, App
Sequence 10910, A
Sequence 1690, Ap
Sequence 1690, Ap
Sequence 1883, Ap

Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 0

US-09-782-953-9 (1-197) x US-09-782-953-8 (1-594)

QY 1 MetProAlaProSerMetAspCyAspValSerThrLeuValAlaCyValValAspVal 20
 Db 1 ATGCCAGCCCTCAGCATGACCTGTGATGTTCCACCTGGTGGCTGTGTGTGATG 60
 QY 21 GluValPheThrAspGlnGluValLysGluLysPheGluGlyLeuPheArgThrTyAsp 40
 Db 61 GAGGTCTTACCAATCAGAGGTTAAGGAAAATTCGAGGACTGTTCCGACCTATGAT 120
 QY 41 GluCyValThrPheGlnLeuPheLysSerPheArgValArgIleAsnPheSerHis 60
 Db 121 GATGTGTGACGTTCCAGCTGTTTAAGACTTTCGACGGGTTGAAATTAATTCAGCCAT 180
 QY 61 PolySerAlaAlaArgAlaArgIleGluLeuHisGluThrGlnPheArgGlyLysLys 80
 Db 181 CCCAATCTGCAGCCCGTCCCGATAGAGCTTCATGAGACTCAGTTCCAGAGGAAAG 240
 QY 81 LeuLysLeuThrPheAlaGlnValGlnThrProGluThrAspGlyAspLysLeuHisLeu 100
 Db 241 CTAAAACTCTACTCTGCCAGGTCCAGACCCAGACAGATGAGACAAACTGCACTTGG 300
 QY 101 AlaProGlnProAlaLysGlnPheLeuLysSerProSerSerProProValGly 120
 Db 301 GACCTCCACACCTGCCAAACAGTTCTCATCTCCACCTTCCTCTCTCTGTGGC 360
 QY 121 TrpLysProIleSerAspAlaThrProValLeuAsnTyraAspLeuLeuTyraValAla 140
 Db 361 TGGAGCCCTATCAGCATGCCACACAGTCCCACTATGACTTCTTATCTGTGGCC 420
 QY 141 LysLeuGlyProGlyGluLysTyrgluLeuHisAlaGlyThrGlnSerThrProSerVal 160
 Db 421 AAACCTAGGACCGAGAGAAATATGAGCTGCACGCTGGAACCTGAGCTACCGAGACCTG 480
 QY 161 ValValHisValCysAspSerAspMetGluGluGluAspProLysThrSerProLys 180
 Db 481 GTGGTGCATGTGTGTACAGCGCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
 QY 181 PolyIleIleGlnThrArgArgProGlyLeuProProSerValSerAsn 197
 Db 541 CCAAAATCATTCAGACCCGGCGTCCGGGCTTGCCACCTCCGTGTCAAC 591

RESULT 2
 US-09-954-456-497
 / Sequence 497, Application US/09954456
 / Patent No. US20020115057A1
 / GENERAL INFORMATION:
 / APPLICANT: Young, Paul
 / TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc
 / FILE REFERENCE: 689290-76
 / CURRENT APPLICATION NUMBER: US/09/954,456
 / PRIOR FILING DATE: 2001-09-18
 / PRIOR APPLICATION NUMBER: US/60/233,617
 / PRIOR FILING DATE: 2000-09-18
 / PRIOR APPLICATION NUMBER: US/60/234,052
 / PRIOR FILING DATE: 2000-09-20
 / PRIOR APPLICATION NUMBER: US/60/234,923
 / PRIOR FILING DATE: 2000-09-25
 / PRIOR APPLICATION NUMBER: US/60/235,134
 / PRIOR FILING DATE: 2000-09-25
 / PRIOR APPLICATION NUMBER: US/60/235,637
 / PRIOR FILING DATE: 2000-09-26
 / PRIOR APPLICATION NUMBER: US/60/235,638
 / PRIOR FILING DATE: 2000-09-26
 / PRIOR APPLICATION NUMBER: US/60/235,711
 / PRIOR FILING DATE: 2000-09-27
 / PRIOR APPLICATION NUMBER: US/60/235,720

PRIOR FILING DATE: 2000-09-27
 / PRIOR APPLICATION NUMBER: US/60/235,840
 / PRIOR FILING DATE: 2000-09-27
 / PRIOR APPLICATION NUMBER: US/60/235,863
 / PRIOR FILING DATE: 2000-09-27
 / NUMBER OF SEQ ID NOS: 2276
 / SOFTWARE: PatentIn version 3.0
 / SEQ ID NO 497
 / LENGTH: 3184
 / TYPE: DNA
 / ORGANISM: Homo sapiens
 US-09-954-456-497

Alignment Scores:
 Pred. No.: 3.77e-102 Length: 3184
 Score: 1007.00 Matches: 190
 Percent Similarity: 98.98% Conservative: 5
 Best Local Similarity: 96.45% Mismatches: 2
 Query Match: 96.73% Indels: 0
 Gaps: 0

US-09-782-953-9 (1-197) x US-09-954-456-497 (1-3184)

QY 1 MetProAlaProSerMetAspCyAspValSerThrLeuValAlaCyValValAspVal 20
 Db 190 ATGCCAGCCCTCAGCATGACCTGTGATGTTCCACCTGGTGGCTGTGTGTGATG 249
 QY 21 GluValPheThrAspGlnGluValLysGluLysPheGluGlyLeuPheArgThrTyAsp 40
 Db 250 GAGGTCTTACCAATCAGAGGTTAAGGAAAATTTGGGGACGTTTCCGACCTTATGAT 309
 QY 41 GluCyValThrPheGlnLeuPheLysSerPheArgValArgIleAsnPheSerHis 60
 Db 310 GACTGTGAGAGTTCCACCTATTAAGAGTTTCAGACCTGTCGATTAATTCAGCAAT 369
 QY 61 PolySerAlaAlaArgAlaArgIleGluLeuHisGluThrGlnPheArgGlyLysLys 80
 Db 370 CTTAAATCTGACGCCCGAGCTAGATGAGCTTCTTGAACCCCAATTCAGAGGAAAATA 429
 QY 81 LeuLysLeuThrPheAlaGlnValGlnThrProGluThrAspGlyAspLysLeuHisLeu 100
 Db 430 TTAAGCTCTACTTGTGCACAGGTTCCAGCTCCAGAGACAGATGAGAGAACTGCACTTG 489
 QY 101 AlaProGlnProAlaLysGlnPheLeuLysSerProSerSerProProValGly 120
 Db 490 GCTCACCCACGCCCGCCAAACAGTTTCTCATCTGCCCTTCCTCCACCTGTTAGC 549
 QY 121 TrpLysProIleSerAspAlaThrProValLeuAsnTyraAspLeuLeuTyraValAla 140
 Db 550 TGGCAGCCCATCAGATGCCACGCACTCTCACTATGACCTCTATAGCTGTGGC 609
 QY 141 LysLeuGlyProGlyGluLysTyrgluLeuHisAlaGlyThrGlnSerThrProSerVal 160
 Db 610 AAACCTAGGACCGAGAGAAATATGAGCTGCACGCTGGAACCTGAGCTACCGAGACCTG 669
 QY 161 ValValHisValCysAspSerAspMetGluGluGluAspProLysThrSerProLys 180
 Db 670 GTGGTGCATGTGTGTACAGCGCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 729
 QY 181 PolyIleIleGlnThrArgArgProGlyLeuProProSerValSerAsn 197
 Db 730 CCAAAATCATTCAAACTGGGCTGTGCTGCTGCCACCTCCGTGTCAAC 780

RESULT 3
 US-09-782-953-17
 / Sequence 17, Application US/09782953
 / Patent No. US20020150953A1
 / GENERAL INFORMATION:
 / APPLICANT: WILLIAMS, R. SANDERS
 / TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
 / FILE REFERENCE: UTSD:674P21


```

; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: US/09/782,953
; CURRENT APPLICATION NUMBER: 2001-02-13
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO: 11
; LENGTH: 2331
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (144)..(734)
US-09-782-953-11

```

```

Alignment Scores:
Pred. No.: 9,54e-62 Length: 2331
Score: 641.00 Matches: 120
Percent Similarity: 79.68% Conservative: 29
Best Local Similarity: 64.17% Mismatches: 30
Query Match: 61.58% Indels: 8
DB: 10 Gaps: 2

```

```

US-09-782-953-9 (1-197) x US-09-782-953-11 (1-2331)
QY 11 SerThrLeuValAlaCysValValAspValGluValPheThrAsnGlnGluValIleGlu 30
DB 174 AGCTCCCGATTGCTGCTGTGGCAACAGTATCTTACGGAAGTAAACCCAGGGCC 233
QY 31 LysPheGluGlyLeuPheArgThrTyraAspGluCysValThrPheGlnLeuPheLysSer 50
DB 234 AAATTGAGTCCCTCTTACGACGTATGCAAGACATACCTTCAGTATTTTAAAGAC 293
QY 51 PheArgValAlaArgIleAsnPheSerHisProLysSerAlaIleArgIleGlu 70
DB 294 TTCAACAGATGAGATTAACCTTCAGCAACCCCTTCTCCGACAGATGCCAGGCTCCAG 353
QY 71 LeuHisGluThrGlnPheArgGlyLysLysLeuLysLeuTyrrPheAlaGlnValGlnThr 90
DB 354 CTGCATTAAGACTGAGTTCTGGGAAAGAAATGAATGATTATTTGCTCAGACCTTACAC 413
QY 91 ProGluThrAspGlyAspLysLeuHisLeuAlaProGlnProAlaLysGlnPheLeu 110
DB 414 ATAGGAAGCTCA-----CACCTGGCTCCGCCAAATCCAGACAAAGCATTTCTG 461
QY 111 IleSerProProSerSerProProValGlyTrpLysProIleSerAspAlaThrProVal 130
DB 469 ATCTCCCTCCCGCTCTCCGCGCAGGTGGATGGAACAAGTGAATGCCAGCCAGTC 521
QY 131 LeuAsnTyraAspLeuLeuTyraValAlaLysLeuGlyProGlyGluLysTyrrGluLeu 150
DB 522 ATAACTATGATCTCTTATATGCCATCTCCAGCTGGGGCCAGGGGAAAGATGATG 581
QY 151 HisAlaGlyThrGlnLysThrProSerValValHisValCysAspSerAspMetGlu 170
DB 582 CACGAGCAGCATGACACCACTCCACGCTGGTGTCTCATGTATGATGATGATCAAGAG 641
QY 171 GluGluGluAspProLysThrSer-----ProLysProLysIleIleGlnThr 186
DB 642 AAGGAGGAAGAAGGAATGAAAGATGAGAGACCTAAGCCAAATTTATCCAGACC 701
QY 187 ArgArgProGlyLeuProPro 193
DB 702 AGGAGGCCGAGATACACGCCG 722

```

RESULT 6
 US-10-044-090-255
 ; Sequence 255, Application US/10044090
 ; Patent No. US20020137081A1

```

; GENERAL INFORMATION:
; APPLICANT: Olga Bandman
; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
; FILE REFERENCE: PA-0028 US
; CURRENT APPLICATION NUMBER: US/10/044,090
; PRIOR FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: PERL Program
; SEQ ID NO: 255
; LENGTH: 2355
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020137081A1 5511889CB1
US-10-044-090-255

```

```

Alignment Scores:
Pred. No.: 9.68e-62 Length: 2355
Score: 641.00 Matches: 120
Percent Similarity: 79.68% Conservative: 29
Best Local Similarity: 64.17% Mismatches: 30
Query Match: 61.58% Indels: 8
DB: 12 Gaps: 2

```

```

US-09-782-953-9 (1-197) x US-10-044-090-255 (1-2355)
QY 11 SerThrLeuValAlaCysValValAspValGluValPheThrAsnGlnGluValIleGlu 30
DB 181 AGCTCCCGATTGCTGCTGTGGCAACAGTATCTTACGGAAGTAAACCCAGGGCC 240
QY 31 LysPheGluGlyLeuPheArgThrTyraAspGluCysValThrPheGlnLeuPheLysSer 50
DB 241 AAATTGAGTCCCTCTTACGACGTATGCAAGACATACCTTCAGTATTTTAAAGAC 300
QY 51 PheArgValAlaArgIleAsnPheSerHisProLysSerAlaIleArgIleGlu 70
DB 301 TTCAACAGATGAGATTAACCTTCAGCAACCCCTTCTCCGACAGATGCCAGGCTCCAG 360
QY 71 LeuHisGluThrGlnPheArgGlyLysLysLeuLysLeuTyrrPheAlaGlnValGlnThr 90
DB 361 CTGCATTAAGACTGAGTTCTGGGAAAGAAATGAATGATTATTTGCTCAGACCTTACAC 420
QY 91 ProGluThrAspGlyAspLysLeuHisLeuAlaProGlnProAlaLysGlnPheLeu 110
DB 421 ATAGGAAGCTCA-----CACCTGGCTCCGCCAAATCCAGACAAAGCATTTCTG 468
QY 111 IleSerProProSerSerProProValGlyTrpLysProIleSerAspAlaThrProVal 130
DB 469 ATCTCCCTCCCGCTCTCCGCGCAGGTGGATGGAACAAGTGAATGCCAGCCAGTC 528
QY 131 LeuAsnTyraAspLeuLeuTyraValAlaLysLeuGlyProGlyGluLysTyrrGluLeu 150
DB 529 ATAACTATGATCTCTTATATGCCATCTCCAGCTGGGGCCAGGGGAAAGATGATG 588
QY 151 HisAlaGlyThrGlnLysThrProSerValValHisValCysAspSerAspMetGlu 170
DB 589 CACGAGCAGCATGACACCACTCCACGCTGGTGTCTCATGTATGATGATGATCAAGAG 648
QY 171 GluGluGluAspProLysThrSer-----ProLysProLysIleIleGlnThr 186
DB 649 AAGGAGGAAGAAGGAATGAAAGATGAGAGACCTAAGCCAAATTTATCCAGACC 708
QY 187 ArgArgProGlyLeuProPro 193
DB 709 AGGAGGCCGAGATACACGCCG 729

```

RESULT 7
 US-09-782-953-23
 ; Sequence 23, Application US/09782953
 ; Patent No. US20020150953A1
 ; GENERAL INFORMATION:
 ; APPLICANT: WILLIAMS, R. SANDERS

APPLICANT: ROTHERMEL, BEVERLY
TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
FILE REFERENCE: US/09/782,953
CURRENT APPLICATION NUMBER: 60/216,601
PRIORITY FILING DATE: 2001-02-13
PRIORITY FILING DATE: 2000-07-07
NUMBER OF SEQ ID NOS: 27
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 23
LENGTH: 720
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (2)..(637)
US-09-782-953-23

Alignment Scores:
Pred. No.: 1,29e-61 Length: 720
Score: 633.50 Matches: 127
Percent Similarity: 77.01% Conservative: 17
Best Local Similarity: 67.91% Mismatches: 32
Query Match: 60.85% Indels: 11
DB: 10 Gaps: 2

US-09-782-953-9 (1-197) x US-09-782-953-23 (1-720)

QY 11 SerThrLeuValAlaCysValValAspValGluValPheThrAsnGlnGluValValGlu 30
DB 83 ACTCAGCTTTTGTCTGCGCGCTCCATGAAGCAGTGTTTTCAGGCGAGAGAGAGAA 142
QY 31 LysPheGluGlyLeuPheArgThrTyrAspGluCysValThrPheGlnLeuPheLysSer 50
DB 143 AGATTTGAAGCACTCTTCCACATCTATGATGACCAAGTTACTTTTCAGCTGTTTAAAGC 202
QY 51 PheArgArgValArgLeuAsnPheSerHisProLysSerAlaAlaArgAlaGlu 70
DB 203 TTTAGAGAGTCAAGATAAATTTTCAGCAAACTTGAAGCGGCGAGAGCGGAGTAA 262
QY 71 LeuHisGluThrGlnPheArgGlyLysLeuLysLeuLysLeuLysLeuLysLeu 90
DB 263 CTCACGAAACACACTTCAATGGCGAGAGCTAAGCTATATTTTGCACAG----- 313
QY 91 ProGluThrAspGlyAspLysLeuHisLeuAlaProGlnProAlaLysGlnPheLeu 110
DB 314 -----TCTATCTCTCGCGCCCGCTGTCAAGCAGTTCTCTC 352
QY 111 IleSerProSerSerProValGlyTyrLysProLysSerAspAlaThrProVal 130
DB 353 ATCTCCCTCCAGCCTCTCCCGCAGTGGGTGGAAGCAGAGATGCGATGCTGTT 412
QY 131 LeuAsnTyrAspLeuLeuTyrAlaValAlaLysLeuGlyProGlyGluLysTyrGluLeu 150
DB 413 ATAAATATGATTTACTCTGTGCTGTTTCCAAATGGGACAGAGAGAAATATGAATT 472
QY 151 HisAlaGlyThrGluSerThrProSerValValHisValCysAspSerAspMetGlu 170
DB 473 CACCGGGAAACAGAGTCGACACCCAGCGTGGTTCATGCTGTGTAAGTGAACCTGAA 532
QY 171 GluGluGluAspProLysThrSerProLysProLysLeuLeuLeuLeuLeuLeu 190
DB 533 GAGGAAGAAGAGACAAAA---AACCCCAACAGAAAAATTGCCCAACAAGGCGCCCGAC 589
QY 191 LeuProProSerValSerAsn 197
DB 590 CTTCCGACCGCAGCTTGAAT 610

RESULT 8

US-09-782-953-5
Sequence 5, Application US/09782953
Patent No. US20020150953A1

GENERAL INFORMATION:
APPLICANT: WILLIAMS, R. SANDERS
APPLICANT: ROTHERMEL, BEVERLY
TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
FILE REFERENCE: US/09/782,953
CURRENT APPLICATION NUMBER: 60/216,601
PRIORITY FILING DATE: 2001-02-13
PRIORITY FILING DATE: 2000-07-07
NUMBER OF SEQ ID NOS: 27
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 5
LENGTH: 597
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(594)
US-09-782-953-5

Alignment Scores:
Pred. No.: 1,28e-61 Length: 597
Score: 632.50 Matches: 121
Percent Similarity: 77.13% Conservative: 24
Best Local Similarity: 64.36% Mismatches: 34
Query Match: 60.76% Indels: 9
DB: 10 Gaps: 2

US-09-782-953-9 (1-197) x US-09-782-953-5 (1-597)

QY 11 SerThrLeuValAlaCysValValAspValGluValPheThrAsnGlnGluValValGlu 30
DB 31 AGCTCCCTGATTTGTTGTGCGCAACGATGATGTTTCAGGCAAGAGTGACAGGGCC 90
QY 31 LysPheGluGlyLeuPheArgThrTyrAspGluCysValThrPheGlnLeuPheLysSer 50
DB 91 AAATTTGAATCTCTTCCAGAACATATGACAGCAGCACCTTCCAGTATTTTAAAGAC 150
QY 51 PheArgArgValArgLeuAsnPheSerHisProLysSerAlaAlaArgAlaGlu 70
DB 151 TTTCAACAGTGTCCGATAACTTCAGCAACCTTATCTGCAGCGGATGCCAGGCTGCGG 210
QY 71 LeuHisGluThrGlnPheArgGlyLysLeuLysLeuLysLeuLysLeuLysLeu 90
DB 211 CTGCACAGCCGAGTCTCTGGGAGGAAATGAAGTTGATTTTCTCAGACTTTACAC 270
QY 91 ProGluThrAspGlyAspLysLeuHisLeuAlaProGlnProAlaLysGlnPheLeu 110
DB 271 ATAGGAAGTTCA-----CACCTGGCTCGCCCAATCCCGACAAACAGTTCTCTC 318
QY 111 IleSerProSerSerProValGlyTyrLysProLysSerAspAlaThrProVal 130
DB 319 ATCTCCCTCCCGCTCTCTCCCTGCTGCGAAACAGTAGAAGATGCCACCCCGCTC 378
QY 131 LeuAsnTyrAspLeuLeuTyrAlaValAlaLysLeuGlyProGlyGluLysTyrGluLeu 150
DB 379 ATAAATATGATTTACTCTGTGCTGTTTCCAAATGGGACAGAGAGAAATATGAATT 438
QY 151 HisAlaGlyThrGluSerThrProSerValValHisValCysAspSerAspMetGlu 170
DB 439 CATGCGGACAGAGCCCACTCCCAAGTGTGTGTCACGTCGTGTGTGTGTGTGTGTGT 498
QY 171 GluGluGluAspProLys-----ThrSerProLysProLysLeuLeuLeuLeu 185
DB 499 AATGAGGAGGAGAGAGAGATGAGAGAAATGAAGAGACCCCAAGCCCAAAATCATCCAG 558
QY 186 ThrArgArgProGlyLeuProPro 193
DB 559 ACACCGGAGCCGAGTACACCG 582

RESULT 9

US-09-782-953-14

Sequence 14, Application US/09782953
 Patent No. US20020150953A1
 GENERAL INFORMATION:
 APPLICANT: WILLIAMS, R. SANDERS
 APPLICANT: ROTHERMEL, BEVERLY
 TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
 FILE REFERENCE: US/09/782,953
 CURRENT FILING DATE: 2001-02-13
 PRIOR APPLICATION NUMBER: 60/216,601
 PRIOR FILING DATE: 2000-07-07
 NUMBER OF SEQ ID NOS: 27
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 14
 LENGTH: 2212
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (125)..(615)
 US-09-782-953-14

Alignment Scores:
 Pred. No.: 3,66-60 Length: 2212
 Score: 626.50 Matches: 121
 Percent Similarity: 74.63% Conservative: 29
 Best Local Similarity: 60.20% Mismatches: 34
 Query Match: 60.18% Indels: 17
 DB: 10 Gaps: 3

US-09-782-953-9 (1-197) x US-09-782-953-14 (1-2212)

QY 6 MetAapCyAspVal-----SerThrLeuValAlaCys 16
 DB 13 ATTGACTGCGAGATGAGAGAGGTGACCTGACGAGACTCCGACCCGACCATCGCTCT 72
 QY 17 ValValAspValGluValPheThrAengInguValLyseGluYpHeGluGluYpHe 36
 DB 73 CACCTGGACCCCGCGGTGTCGCGACGCGCTGCGGCAATTGAGTCCCTCTT 132
 QY 37 ArgThrTyraPglucYsValThrPheGlnLeuPheLysSerPheArgValAlaGlie 56
 DB 133 AGGACCTATGACAGAGACATCACTTCAGTATTMAAGCTTCAACGAGTCAGATA 192
 QY 57 AnpPheSerHisProLysSerAlaAlaArgAlaArgIleGluLeuHisGluThrGlnPhe 76
 DB 193 AACTTCAGCAACCCCTTCCTCCGACGAGATGCCAGCTCCAGCTGCATAGACTGATT 252
 QY 77 ArgGlyLyLeuLeuLeuYrPheAlaGlnValGlnThrProGluThrAspGlyAsp 96
 DB 253 CCGGGAAGGAATGAAATGATATTTGCTCAGACTTACATAGAGACTCA----- 306
 QY 97 LybLeuHisLeuAlaProProGlnProAlaLybGlnPheLeuLysSerProPheSer 116
 DB 307 -----CACCTGGCTCCGCAATCCAGACAGACATTTCTGATCTCCCTCCGCTCT 360
 QY 117 ProProValGlyTyraPglucYsValThrPheAlaGlnValGlnThrProGluThr 136
 DB 361 CCGCCGCTGGATGAGAAACAGTGAAGATGCAACCCGACCTCAATACATATGATCTCTTA 420
 QY 421 TATGCATCTCCAACTGGGGCCAGGGAAGATGATGATGCAACGAGGATGACACCC 480
 QY 137 TyraValAlaAlaLybLeuGlyProGlyGluLyTyrgLuleuHisAlaGlyThrGlnSer 156
 DB 421 TATGCATCTCCAACTGGGGCCAGGGAAGATGATGATGCAACGAGGATGACACCC 480
 QY 157 ThrProSerValValAlaHisValIleCyAspSerAspMetGluGluGluAspProLy 176
 DB 481 ACTCCAGCGGTGTGTCTCATGTATGTGAGAGTCAAGAGAGAGAGAGAGAGAGAA 540
 QY 177 ThrSer-----ProLybProLybIleIleGlnThrArgArgProGlyLeuPro 192
 DB 541 ATGGAAAGATGAGAGACTTAAGCCAAAATTTATCCAGCCAGGAGGCGGAGTACAG 600

QY 193 Pro 193
 DB 601 CCG 603

RESULT 10

US-09-782-953-2
 Sequence 2, Application US/09782953
 Patent No. US20020150953A1
 GENERAL INFORMATION:
 APPLICANT: WILLIAMS, R. SANDERS
 APPLICANT: ROTHERMEL, BEVERLY
 TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
 FILE REFERENCE: US/09/782,953
 CURRENT FILING DATE: 2001-02-13
 PRIOR APPLICATION NUMBER: 60/216,601
 PRIOR FILING DATE: 2000-07-07
 NUMBER OF SEQ ID NOS: 27
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 2
 LENGTH: 597
 TYPE: DNA
 ORGANISM: Mus musculus
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (1)..(594)
 US-09-782-953-2

Alignment Scores:
 Pred. No.: 1,64e-59 Length: 597
 Score: 613.50 Matches: 119
 Percent Similarity: 75.00% Conservative: 22
 Best Local Similarity: 63.30% Mismatches: 38
 Query Match: 58.93% Indels: 9
 DB: 10 Gaps: 2

US-09-782-953-9 (1-197) x US-09-782-953-2 (1-597)

QY 11 SerThrLeuValAlaCysValValAspValGluValPheThrAengInguValLyseGlu 30
 DB 31 AGCGGACCATCGCTGCTGACCTGACGAGCCGCGGTGTCGCGAGCGGCTGCGGCGCC 90
 QY 31 LybPheGluGlyLeuPheArgThrTyraPglucYsValThrPheGlnLeuPheLysSer 50
 DB 91 AAATTGAAATCCCTCTTCAGAAATATGACAAAGACACACCTTCAGTATTMAAGAC 150
 QY 51 PheArgValAlaArgIleAspPheSerHisProLysSerAlaAlaArgAlaGlieGlu 70
 DB 151 TTCAACGCTGCGGATTAACCTTCAGCAACCCCTTATCTGACCGGATGCCAGGCTGCGG 210
 QY 71 LeuHisGluThrGlnPheArgGlyLybLybLeuYrPheAlaGlnValGlnThr 90
 DB 211 CTGCAACAGACCGAATTCCTGCGGGAAGAAATGATGATTTGCTCAGACTTACAC 270
 QY 91 ProGluThrAspGlyAspLybLeuHisLeuAlaProProGlnProAlaLybGlnPheLeu 110
 DB 271 ATAGGAAGTTCA-----CACCTGGCTCCGCAATCCGACCAACAGATTCTCTC 318
 QY 111 IleserProPheSerProProValGlyTyraPglucYsValThrPheAlaGlnThr 130
 DB 319 ATCTCCCTCGGCTCTCTCTCCGTTGCGGAAACAGTGAAGATGCAACCCCGCTC 378
 QY 131 LeuAsnTyraAspLeuLeuTyraValAlaLybLeuGlyProGlyGluLyTyrgLuleu 150
 DB 379 ATAAATTAGATCTTTATATCTCATCTCCAAAGCTGGGGCCAGGAGAGATGACG 438
 QY 151 HisAlaGlyThrGlnSerThrProSerValValAlaHisValIleCyAspSerAspMetGlu 170
 DB 439 CATGACGACGACGACCCCACTCCAGTGTGTGTCCACGTTGTGAGAGTCAACAGAG 498
 QY 171 GluGluGluAspProLyb-----ThrSerProLybProLybIleIleGln 185

Db 499 AATGAGGAGGAGGAGGAGATCGAGAGATGAAGAGACCAAGCCCAAAATCATCCAG 558
Qy 186 ThrArgArgProGlyLeuProPro 193
Db 559 ACACGAGACCGGAGTACACACG 582
RESULT 11
US-09-880-107-3340
; Sequence 3340, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3340
; LENGTH: 2173
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 U28833
US-09-880-107-3340
Alignment Scores:
Pred. No.: 2,36e-57 Length: 2173
Score: 601.00 Matches: 115
Percent Similarity: 79.77% Conservative: 23
Best Local Similarity: 66.47% Mismatches: 27
Query Match: 57.73% Indels: 8
DB: 10 Gaps: 2
US-09-782-953-9 (1-197) x US-09-880-107-3340 (1-2173)
Qy 25 AsnGlnGluValLysGluLysPheGluGlyLeuPheArgThrTyrAspGluCysValThr 44
Db 43 AACAGATGGTGTATGCCAAATTTGAGTCCCTCTTTAGGACGTATGACAGGACATCACC 102
Qy 45 PheGlnLeuPheLysSerPheArgValArgLeuPheGlnPheArgGlyLysLysLeuLysLeuTyr 84
Db 103 TTTGAGTATTTTAAAGCTTCAACAGAGTCAAGATTAACCTTCAGCAACCTTCTCCGCA 162
Qy 65 AlaArgAlaArgLeuGluLeuHisGluThrGlnPheArgGlyLysLysLeuLysLeuTyr 84
Db 163 GCAGATGCCAGGCTCCAGCTGCATAGACGTAGTTCTGCGAAGGAAATGAAGTATAT 222
Qy 85 PheAlaGlnValGlnThrProGluThrAspGlyAspLysLeuHisLeuAlaProProGln 104
Db 223 TTTGCTCAGACCTTACACATAGGAAGCTCA-----CACCTGGCTCCGCCAAAT 270
Qy 105 ProAlaLysGlnPheLeuLysSerProProSerSerProValGlyTyrLysProLeu 124
Db 271 CCAGACAGGAGTTCTCGATCTCCCTCCGCTCTCCGCGAGTGGATGGAAACAGTG 330
Qy 125 SerAspAlaThrProValLeuAsnTyrAspLeuLeuTyrAlaValAlaLysLeuGlyPro 144
Db 331 GAAGATCGGACCCAGTCATAAATATGATCTCTTATATGCCATCTCCAGAGTGGGGCCA 390
Qy 145 GlyGluLysTyrGluLeuHisAlaGlyThrGluSerThrProSerValValValHisVal 164
Db 391 GGGGAAAGTATGAATTTGACGCGAGCGACTGACACCACTCCAGCGTGGTGGTCCATGTA 450
Qy 165 CysAspSerAspMetGluGluGluAspProLysThrSer-----ProLys 180
Db 165 CysAspSerAspMetGluGluGluAspProLysThrSer-----ProLys 180

Db 451 TGTGAGAGTGTATCAAGAGAGGAGGAGGAGAAATGGAAGAATGAGGAGACCTAAG 510
Qy 181 ProLysIleLeuGlnThrArgArgProGlyLeuProPro 193
Db 511 CCAAAATTTATCCAGACCGAGGAGCGCGAGTACACCGCG 549
RESULT 12
US-09-925-302-347
; Sequence 347, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 347
; LENGTH: 2358
; TYPE: DNA
; ORGANISM: Homo sapiens
; OTHER INFORMATION: US-09-925-302-347
Alignment Scores:
Pred. No.: 1.79e-55 Length: 2358
Score: 594.50 Matches: 119
Percent Similarity: 73.13% Conservative: 28
Best Local Similarity: 59.20% Mismatches: 37
Query Match: 56.15% Indels: 18
DB: 10 Gaps: 3
US-09-782-953-9 (1-197) x US-09-925-302-347 (1-2358)
Qy 6 MetAspCysAspVal-----SerThrLeuValAlaCys 16
Db 102 ATTGACTGGAGATGGAGGAGGTGGACCTGCAGGACCTGCCAGCCACCATCGCTGT 161
Qy 17 ValValAspValGluValPheThrAsnGlnGluValLysGluLysPheGluGlyLeuPhe 36
Db 162 CACCTGGACCCCGCGTGTCTGGACGCGCTGTGCGGCGCCAAATTTGAGTCCCTCTT 221
Qy 37 ArgThrTyrAspGluCysValThrPheGlnLeuPheLysSerPheArgValArgIle 56
Db 222 AGGACGTATGACAAGGACATCACCTTTCAGTATTTTAAAGCTTCAACGAGTCAGAATA 281
Qy 57 AsnPheSerHisProLysSerAlaAlaArgAlaArgIleGluLeuHisGluThrGlnPhe 76
Db 282 AACTTCAGCAACCCCTTCTCCGACGAGATGCGAGCTCCAGCTGCATAGACTGAGTTT 341
Qy 77 ArgGlyLysLeuLysLeuTyrPheAlaGlnValGlnThrProGluThrAspGlyAsp 96
Db 342 CTGGGAAAGAAATGAAGTTATATTTTGTCTCAGACCTTACACATAGGAAGCTCA----- 395
Qy 97 LysLeuHisLeuAlaProProGlnProAlaLysGlnPheLeuLysSerProProSer 116
Db 396 -----CACCTGGCTCCGCA-AATCCAGACAAGCAGTTTCTGATCTCCCGCTCT 448
Qy 117 ProProValGlyTyrLysProLeuSerAspAlaThrProValLeuAsnTyrAspLeu 136
Db 449 CCGCAGTGGGATGGAAACAGTGGAAAGATGCGACCCCAAGTCAATAACTATATCTCTTA 508
Qy 137 TyrAlaValAlaLysLeuGlyProGlyGluLysTyrGluLeuHisAlaGlyThrGluSer 156
Db 509 TATGCCATCTCCAAGCTGGGGCGAGGGGAAAGTATGAATTCACGCGACTGACACC 568
Qy 157 ThrProSerValValHisValCysAspSerAspMetGluGluGluAspProLys 176
Db 569 ACTCCAGCGTGTGCTCCATGTATGTAGAGTGTATGATCAAGAGAGGAGGAGGAGAA 628

Qy 177 Thrsr-----ProlyseProlyleleleInThrArgProGlyLeuPro 192
Db 629 ATGAAAGATGAGAGACCTTAAGCAAAAATTATCCAGACGAGCGCGGATGACAG 688
Qy 193 Pro 193
Db 689 CCG 691
RESULT 13
US-09-782-953-1
Sequence 1, Application US/09782953
Patent No. US20020150953A1
GENERAL INFORMATION:
APPLICANT: WILLIAMS, R. SANDERS
APPLICANT: ROTHERMEL, BEVERLY
TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
FILE REFERENCE: US/09-782,953
CURRENT FILING DATE: 2001-02-13
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/216,601
NUMBER OF SEQ ID NOS: 27
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 1
LENGTH: 599
TYPE: DNA
ORGANISM: Mus musculus
US-09-782-953-1

Alignment Scores:
Pred. No.: 5,1e-54 Length: 599
Score: 564.00 Matches: 109
Percent Similarity: 81.10% Conservative: 24
Best Local Similarity: 66.46% Mismatches: 27
Query Match: 54.18% Indels: 5
DB: 10 Gaps: 1
US-09-782-953-9 (1-197) x US-09-782-953-1 (1-599)

Qy 11 SerThrLeuValAlaCysValValAlaPheValGluValPheThrArgInGluValValysGlu 30
Db 110 ACCTCCCTGATGCTGTGTGGCAACGATGATGCTTACGCAAGAGTACAGACGAGCGC 169
Qy 31 LysPheGluGlyLeuPheArgThrTyraPgluCyValThrPheGlnLeuPheLysSer 50
Db 170 AATTTGAATCCCTCTTCAGACATATGACACAGACACACCTTCCAGATATTTTAAGAC 229
Qy 51 PheArgValAlaArgIleAsnPheserH1ProlyseR1AlaAlaArgIleGlu 70
Db 230 TTCAAACGATCGCGATTAACCTCACCAACCCCTTATTCGACGCGATGCGCGCG 289
Qy 71 LeuH1sgluthrGlnPheArgGlyLysLysLeuLysLeuTyraPheAlaGlnValGlnThr 90
Db 290 CTGCACAAAGACCCAGTTCCTGCGGAGGAATCAAGTTGATTTCTCAGACTTTACAC 349
Qy 91 ProGluThrAspGlyAspLysLeuH1sleuAlaProGlnProAlaLysGlnPheLeu 110
Db 350 ATAGCAAGTTCA-----CACCTGGCTCCGCCAAT-CCGACAAACAGTTCTC 396
Qy 111 IleserProPserSerProProValGlyTrpLysProIleserAspAlaThrProVal 130
Db 397 ATCTCCCTCCGCGCTCTCTCTCCGTTGGTGAAACAAGTAAAGATGCGACCCCGTC 456
Qy 131 LeuAsnTyraPheLeuLeuTyraValAlaLysLeuGlyProGlyGluLysTyrgluLeu 150
Db 457 ATAAATTAAGATCTTTATATGCAATCTCCAAAGCTGGGCGCAGAGAGAAATGTGAAC 516
Qy 151 HisAlaGlyThrGluSerThrProSerValValHisValLysAspSerAspMetGlu 170
Db 517 CATGACGACGACACACCACTCCAGTGTGTGTCCACGTGTGTAGATGACCAAGAG 576

Qy 171 GluGluGluAsp 174
Db 577 AATGAGAGGANA 588

RESULT 14

US-09-864-761-10388/C
Sequence 10388, Application US/09864761
Patent No. US20020048763A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.

APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecomica-X-1
GENE EXPRESSION ANALYSIS BY MICROARRAY

CURRENT FILING DATE: US/09/864,761

PRIOR FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312

PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: US 60/207,456

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/632,366

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: GB 24263,6

PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/236,359

PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: PCT/US01/00666

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00667

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00664

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00669

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00665

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00668

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00663

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00662

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00661

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00670

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: US 60/234,687

PRIOR FILING DATE: 2000-09-21

PRIOR APPLICATION NUMBER: US 09/608,408

PRIOR FILING DATE: 2000-06-30

PRIOR APPLICATION NUMBER: US 09/774,203

PRIOR FILING DATE: 2001-01-29

NUMBER OF SEQ ID NOS: 49117

SOFTWARE: Annomax Sequence Listing Engine vers. 1.1

SEQ ID NO 10388

LENGTH: 412

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: MAP TO AP000054.1

OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.6

OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.96

OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4

OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.67

OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 1

OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.62

OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.67

US-09-864-761-10388

Alignment Scores:

Pred. No.: 2e-16 Length: 412

Score: 223.00 Matches: 42
Percent Similarity: 89.66% Conservative: 10
Best Local Similarity: 72.41% Mismatches: 6
Query Match: 21.42% Indels: 0
DB: 10 Gaps: 0

US-09-782-953-9 (1-197) x US-09-864-761-10388 (1-412)

Qy 31 LysPheGluGlyLeuPheArgThrTyArgGluCysValThrPheGlnLeuPheLysSer 50
Db 407 AAATTTGAGTCCCTTTAGGACGTATGACAGGACATCACCTTTTCAGTATTTTAAGAGC 348
Qy 51 PheArgValArgIleAsnPheSerHisProLysSerAlaAlaArgAlaGlu 70
Db 347 TTCAACGAGTCAGATAAACTTTCAGAACCCCTTCCGACGAGATGCCAGGCTCCAG 288
Qy 71 LeuHisGluThrGlnPheArgGlyLysLeuLysLeuTy-PheAlaGlnVal 88
Db 287 CTGCATAAGACTGAGTTCTGGGAAAGGAATGAGTTATATTTTCTCAGGTG 234

RESULT 15

US-09-864-761-2064/c
; Sequence 2064, Application US/09864761
; Patent No. US2002048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aomicra-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIORITY FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 2064

LENGTH: 446

TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:

OTHER INFORMATION: MAP TO AF000122.1
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.1
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 4.1
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.7
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.3
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.2
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.4

US-09-864-761-2064

Alignment Scores:

Pred. No.: 2.23e-16 Length: 446
Score: 223.00 Matches: 42
Percent Similarity: 89.66% Conservative: 10
Best Local Similarity: 72.41% Mismatches: 6
Query Match: 21.42% Indels: 0
DB: 10 Gaps: 0

US-09-782-953-9 (1-197) x US-09-864-761-2064 (1-446)

Qy 31 LysPheGluGlyLeuPheArgThrTyArgGluCysValThrPheGlnLeuPheLysSer 50
Db 441 AAATTTGAGTCCCTTTAGGACGTATGACAGGACATCACCTTTTCAGTATTTTAAGAGC 382
Qy 51 PheArgValArgIleAsnPheSerHisProLysSerAlaAlaArgAlaGlu 70
Db 381 TTCAACGAGTCAGATAAACTTTCAGAACCCCTTCCGACGAGATGCCAGGCTCCAG 322
Qy 71 LeuHisGluThrGlnPheArgGlyLysLysLeuTy-PheAlaGlnVal 88
Db 321 CTGCATAAGACTGAGTTCTGGGAAAGGAATGAGTTATATTTTCTCAGGTG 268

Search completed: December 15, 2002, 00:06:30
Job time : 37.8887 secs

THIS PAGE BLANK (USPTO)

```
RESULT 2
US-09-782-953-10
; Sequence 10, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: US/09/782,953
; CURRENT APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2001-02-13
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-782-953-10

Query Match      100.0%; Score 1041; DB 10; Length 197;
Best Local Similarity 100.0%; Pred. No. 1.2e-94;
Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPAPSMDCDVSTLVACVVDVEFTNQEVKEFGLFRTYDECVTQFLFKSFRRVRINFESH 60
DB 1 MPAPSMDCDVSTLVACVVDVEFTNQEVKEFGLFRTYDECVTQFLFKSFRRVRINFESH 60
QY 61 PKSAARARIELHETQFRGKKLYFAQVQTPETDGDGKLHLAPPQAKOFLISPPSSPPVWGKPS 120
DB 61 PKSAARARIELHETQFRGKKLYFAQVQTPETDGDGKLHLAPPQAKOFLISPPSSPPVWGKPS 120
QY 121 WKPISDATPVLYNDLLYAVAKLPGKEKELHAGTSTPSVVHVCDSDMEEDPKTSKPKEIIQ 180
DB 121 WKPISDATPVLYNDLLYAVAKLPGKEKELHAGTSTPSVVHVCDSDMEEDPKTSKPKEIIQ 180
QY 181 PKIITRRPGLPPSVSN 197
DB 181 PKIITRRPGLPPSVSN 197

RESULT 3
US-09-782-953-18
; Sequence 18, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: US/09/782,953
; CURRENT APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2001-02-13
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-18

Query Match      94.1%; Score 980; DB 10; Length 192;
Best Local Similarity 96.4%; Pred. No. 1e-88;
Matches 185; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 6 MDCDVSTLVACVVDVEFTNQEVKEFGLFRTYDECVTQFLFKSFRRVRINFESHPSKAA 65
DB 1 MDCDVSTLVACVVDVEFTNQEVKEFGLFRTYDECVTQFLFKSFRRVRINFESHPSKAA 60
QY 66 RARIELHETQFRGKKLYFAQVQTPETDGDGKLHLAPPQAKOFLISPPSSPPVWGKPS 125
DB 61 RARIELHETQFRGKKLYFAQVQTPETDGDGKLHLAPPQAKOFLISPPSSPPVWGKPS 120
QY 126 DATPVLYNDLLYAVAKLPGKEKELHAGTSTPSVVHVCDSDMEEDPKTSKPKEIIQ 185
DB 121 DATPVLYNDLLYAVAKLPGKEKELHAGTSTPSVVHVCDSDMEEDPKTSKPKEIIQ 180
QY 186 TRRPGLPPSVSN 197
DB 181 TRRPGLPPSVSN 192

RESULT 4
US-09-782-953-19
; Sequence 19, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: US/09/782,953
; CURRENT APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2001-02-13
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-19

Query Match      94.1%; Score 980; DB 10; Length 192;
Best Local Similarity 96.4%; Pred. No. 1e-88;
Matches 185; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 6 MDCDVSTLVACVVDVEFTNQEVKEFGLFRTYDECVTQFLFKSFRRVRINFESHPSKAA 65
DB 1 MDCDVSTLVACVVDVEFTNQEVKEFGLFRTYDECVTQFLFKSFRRVRINFESHPSKAA 60
QY 66 RARIELHETQFRGKKLYFAQVQTPETDGDGKLHLAPPQAKOFLISPPSSPPVWGKPS 125
DB 61 RARIELHETQFRGKKLYFAQVQTPETDGDGKLHLAPPQAKOFLISPPSSPPVWGKPS 120
QY 126 DATPVLYNDLLYAVAKLPGKEKELHAGTSTPSVVHVCDSDMEEDPKTSKPKEIIQ 185
DB 121 DATPVLYNDLLYAVAKLPGKEKELHAGTSTPSVVHVCDSDMEEDPKTSKPKEIIQ 180
QY 186 TRRPGLPPSVSN 197
DB 181 TRRPGLPPSVSN 192

RESULT 5
US-09-782-953-21
; Sequence 21, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: US/09/782,953
; CURRENT APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2001-02-13
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-21

Query Match      94.1%; Score 980; DB 10; Length 192;
Best Local Similarity 96.4%; Pred. No. 1e-88;
Matches 185; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 6 MDCDVSTLVACVVDVEFTNQEVKEFGLFRTYDECVTQFLFKSFRRVRINFESHPSKAA 65
DB 1 MDCDVSTLVACVVDVEFTNQEVKEFGLFRTYDECVTQFLFKSFRRVRINFESHPSKAA 60
QY 66 RARIELHETQFRGKKLYFAQVQTPETDGDGKLHLAPPQAKOFLISPPSSPPVWGKPS 125
DB 61 RARIELHETQFRGKKLYFAQVQTPETDGDGKLHLAPPQAKOFLISPPSSPPVWGKPS 120
QY 126 DATPVLYNDLLYAVAKLPGKEKELHAGTSTPSVVHVCDSDMEEDPKTSKPKEIIQ 185
DB 121 DATPVLYNDLLYAVAKLPGKEKELHAGTSTPSVVHVCDSDMEEDPKTSKPKEIIQ 180
QY 186 TRRPGLPPSVSN 197
DB 181 TRRPGLPPSVSN 192
```

```
QY 66 RARIELHETQFRGKKLYFAQVQTPETDGDGKLHLAPPQAKOFLISPPSSPPVWGKPS 125
DB 61 RARIELHETQFRGKKLYFAQVQTPETDGDGKLHLAPPQAKOFLISPPSSPPVWGKPS 120
QY 126 DATPVLYNDLLYAVAKLPGKEKELHAGTSTPSVVHVCDSDMEEDPKTSKPKEIIQ 185
DB 121 DATPVLYNDLLYAVAKLPGKEKELHAGTSTPSVVHVCDSDMEEDPKTSKPKEIIQ 180
QY 186 TRRPGLPPSVSN 197
DB 181 TRRPGLPPSVSN 192

RESULT 4
US-09-782-953-19
; Sequence 19, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: US/09/782,953
; CURRENT APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2001-02-13
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-19

Query Match      94.1%; Score 980; DB 10; Length 192;
Best Local Similarity 96.4%; Pred. No. 1e-88;
Matches 185; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 6 MDCDVSTLVACVVDVEFTNQEVKEFGLFRTYDECVTQFLFKSFRRVRINFESHPSKAA 65
DB 1 MDCDVSTLVACVVDVEFTNQEVKEFGLFRTYDECVTQFLFKSFRRVRINFESHPSKAA 60
QY 66 RARIELHETQFRGKKLYFAQVQTPETDGDGKLHLAPPQAKOFLISPPSSPPVWGKPS 125
DB 61 RARIELHETQFRGKKLYFAQVQTPETDGDGKLHLAPPQAKOFLISPPSSPPVWGKPS 120
QY 126 DATPVLYNDLLYAVAKLPGKEKELHAGTSTPSVVHVCDSDMEEDPKTSKPKEIIQ 185
DB 121 DATPVLYNDLLYAVAKLPGKEKELHAGTSTPSVVHVCDSDMEEDPKTSKPKEIIQ 180
QY 186 TRRPGLPPSVSN 197
DB 181 TRRPGLPPSVSN 192

RESULT 5
US-09-782-953-21
; Sequence 21, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: US/09/782,953
; CURRENT APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2001-02-13
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-21

Query Match      94.1%; Score 980; DB 10; Length 192;
Best Local Similarity 96.4%; Pred. No. 1e-88;
Matches 185; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 6 MDCDVSTLVACVVDVEFTNQEVKEFGLFRTYDECVTQFLFKSFRRVRINFESHPSKAA 65
DB 1 MDCDVSTLVACVVDVEFTNQEVKEFGLFRTYDECVTQFLFKSFRRVRINFESHPSKAA 60
QY 66 RARIELHETQFRGKKLYFAQVQTPETDGDGKLHLAPPQAKOFLISPPSSPPVWGKPS 125
DB 61 RARIELHETQFRGKKLYFAQVQTPETDGDGKLHLAPPQAKOFLISPPSSPPVWGKPS 120
QY 126 DATPVLYNDLLYAVAKLPGKEKELHAGTSTPSVVHVCDSDMEEDPKTSKPKEIIQ 185
DB 121 DATPVLYNDLLYAVAKLPGKEKELHAGTSTPSVVHVCDSDMEEDPKTSKPKEIIQ 180
QY 186 TRRPGLPPSVSN 197
DB 181 TRRPGLPPSVSN 192
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 11:39:49 ; Search time 6.31197 Seconds
(without Alignments)

506.931 Million cell updates/sec

Title: US-09-782-953-9

Perfect score: 1041

Sequence: 1 MPAPSMDCDVSTLVACVVDV.....SPFKIQRRLPGLPPSVSN 197

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 103943 seqs, 16242309 residues

Total number of hits satisfying chosen parameters: 103943

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Published Applications_Aa.*
2: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB pep.*
8: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep.*
10: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1041	100.0	197	US-09-782-953-9	Sequence 9, Appl
2	1041	100.0	197	US-09-782-953-10	Sequence 10, Appl
3	980	94.1	192	US-09-782-953-18	Sequence 18, Appl
4	980	94.1	192	US-09-782-953-19	Sequence 19, Appl
5	658.5	63.3	241	US-09-782-953-21	Sequence 21, Appl
6	658.5	63.3	241	US-09-782-953-22	Sequence 22, Appl
7	641	61.6	197	US-09-782-953-12	Sequence 12, Appl
8	641	61.6	197	US-09-782-953-13	Sequence 13, Appl
9	633.5	60.9	212	US-09-782-953-24	Sequence 24, Appl
10	633.5	60.9	212	US-09-782-953-25	Sequence 25, Appl
11	632.5	60.8	198	US-09-782-953-6	Sequence 6, Appl
12	632.5	60.8	198	US-09-782-953-7	Sequence 7, Appl
13	622	59.8	197	US-09-782-953-15	Sequence 15, Appl
14	622	59.8	197	US-09-782-953-16	Sequence 16, Appl
15	613.5	58.9	198	US-09-782-953-3	Sequence 3, Appl
16	613.5	58.9	198	US-09-782-953-4	Sequence 4, Appl
17	254.5	24.4	142	US-09-925-302-790	Sequence 790, App
18	219	21.0	58	US-09-864-761-35379	Sequence 35379, A
19	219	21.0	58	US-09-864-761-43076	Sequence 43076, A

20	174	16.7	56	10	US-09-864-761-34111	Sequence 34111, A
21	90	8.6	526	10	US-09-801-368-362	Sequence 362, App
22	85	8.2	4019	10	US-09-738-973-425	Sequence 425, App
23	82	7.9	2783	10	US-09-816-669A-14	Sequence 14, Appl
24	81	7.8	559	10	US-09-854-549-7	Sequence 7, Appl
25	81	7.8	559	10	US-09-854-549-7	Sequence 6, Appl
26	79	7.6	985	9	US-09-978-285A-211	Sequence 211, App
27	79	7.6	985	9	US-09-978-697-211	Sequence 211, App
28	79	7.6	985	9	US-09-978-192A-211	Sequence 211, App
29	78.5	7.5	932	10	US-09-754-997A-2	Sequence 2, Appl
30	78.5	7.5	1252	10	US-09-754-997A-2	Sequence 11, Appl
31	77	7.3	135	10	US-09-858-664A-11	Sequence 2, Appl
32	76.5	7.3	330	10	US-09-794-591-2	Sequence 2, Appl
33	75	7.2	976	10	US-09-969-528-2	Sequence 2, Appl
34	75	7.2	2441	12	US-10-109-886-8	Sequence 8, Appl
35	74	7.1	433	10	US-09-925-300-1467	Sequence 1467, App
36	73	7.0	228	10	US-09-864-761-37276	Sequence 977, App
37	72.5	7.0	282	10	US-09-864-761-35386	Sequence 35386, A
38	72.5	7.0	282	10	US-09-864-761-35386	Sequence 37704, A
39	72.5	7.0	598	10	US-09-853-386-91	Sequence 91, Appl
40	72.5	7.0	598	10	US-09-853-386-91	Sequence 98, Appl
41	72.5	7.0	598	10	US-09-853-386-98	Sequence 43244, A
42	72.5	7.0	748	10	US-09-864-761-43244	Sequence 53, Appl
43	72	6.9	662	10	US-09-799-777-53	Sequence 164, App
44	72	6.9	743	10	US-09-771-161A-164	Sequence 254, App
45	72	6.9	743	10	US-09-771-161A-254	

ALIGNMENTS

```
RESULT 1
US-09-782-953-9
; Sequence 9, Application US/09782953
; Patent No. US20020150953A1
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674P21
; CURRENT APPLICATION NUMBER: US/09/782,953
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-782-953-9

Query Match      100.0%; Score 1041; DB 10; Length 197;
Best Local Similarity 100.0%; Pred. No. 1.2e-94;
Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPAPSMDCDVSTLVACVVDVFTNOEVEKEFGLFRITDECVTFLPFRFRVINFSH 60
DB 1 MPAPSMDCDVSTLVACVVDVFTNOEVEKEFGLFRITDECVTFLPFRFRVINFSH 60
QY 61 PKSAARIEIHEHTQFRGKKLKLYFAOVQTPETDGLHAPQPAKOFILSPSSPPVG 120
DB 61 PKSAARIEIHEHTQFRGKKLKLYFAOVQTPETDGLHAPQPAKOFILSPSSPPVG 120
QY 121 WKISDATPTLVNDLYAAYAKLGPGEKYEIHAHTESTPSVVVAVCDSDMEEDPKTSK 180
DB 121 WKISDATPTLVNDLYAAYAKLGPGEKYEIHAHTESTPSVVVAVCDSDMEEDPKTSK 180
QY 181 PKIQRRLPGLPPSVSN 197
DB 181 PKIQRRLPGLPPSVSN 197
```

```
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Parentin Ver. 2.1
; SEQ ID NO 21
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-953-21

Query Match      63.3%; Score 658.5; DB 21; Length 241;
Best Local Similarity 70.1%; Pred. No. 3.1e-58;
Matches 131; Conservative 17; Mismatches 38; Indels 1; Gaps 1;

Qy 11 STLACVVDEVEFTNQEVKEKEGEGFRYDECVTQLPKSRRVRINFSPKSAARARIE 70
Db 47 TSLFACSVHEAVFEAREOKERPEALFTYDDQVTFQLPKSRRVRINFSPKSAARARIE 106
Qy 71 LHETQFRGKCLKLYPAQVQTPETDGDGLHLAPPQPAKQFLISPPSSPPVGVWKPISDATPV 130
Db 107 LHETDFNGKCLKLYPAQVQMSGEVRDKSYLLPPQPVKQFLISPPASPPVGVWQKQSEDAMPV 166
Qy 131 LNYDLLYAVAKLPGCKYELHAGTSTPSVVVHVCDSDMEDEEDPKTSPKPKIQTTRPG 190
Db 167 INYDLLCAVSKLPGCKYELHAGTSTPSVVVHVCDSETEEEETK-NPKQKIAQTRRPD 225
Qy 191 LPPSVSN 197
Db 226 PPTAALN 232
```

Search completed: December 11, 2002, 11:49:34
Job time : 135.874 secs

Db 249 LPPSVSN 255

RESULT 12

US-09-575-580B-5
; Sequence 5, Application US/09575580B
; GENERAL INFORMATION:
; APPLICANT: McKee, K.
; APPLICANT: McKee, K.
; APPLICANT: Ryeon, S.
; TITLE OF INVENTION: CALCIPRESSINS: ENDOGENOUS INHIBITORS OF CALCINEURIN
; TITLE OF INVENTION: USES AND REAGENTS RELATED THERETO
; FILE REFERENCE: HMV-048.01
; CURRENT APPLICATION NUMBER: US/09/575,580B
; CURRENT FILING DATE: 2000-05-22
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 242
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-575-580B-5

Query Match

Best Local Similarity 84.7%; Score 682; DB 19; Length 242;
Best Local Similarity 91.8%; Pred. No. 4,6e-81;
Matches 168; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

QY 11 STLAVCVDEVEFTNOEVEKEFGLFRTYDECVTFOLFKSFRVRINFSHPKSAARARIE 70

Db 57 NSLFACVHOSVFEESKEKEFGLFRTYDECVTFOLFKSFRVRINFSHPKSAARARIE 116

QY 71 LHETORFGKKLKYFAOVQTPETDGDKLHAPQAKOFLISPPSPVGMKPISDATPV 130

Db 117 LHETORFGKKLKYFAOVQTPETDGDKLHAPQAKOFLISPPSPVGMKPISDATPV 176

QY 131 LNYDLVAVAKLGPGEKYEELHAGTSTPSVVHVCDSDMEEDDPKTSPPKXIQTTRPG 190

Db 177 LNYDLVAVAKLGPGEKYEELHAGTSTPSVVHVCDSDMEEDDPKTSPPKXIQTTRPG 236

QY 191 LPP 193

Db 237 LPP 239

RESULT 13

PCT-US02-17382-133
; Sequence 133, Application PC/TUS0217382
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: MODIFIERS OF THE P53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-062
; CURRENT APPLICATION NUMBER: PCT/US02/17382
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: US 60/296,076
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,605
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/357,253
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 234
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 133
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-17382-133

Query Match

Best Local Similarity 63.3%; Score 658.5; DB 1; Length 241;
Best Local Similarity 70.1%; Pred. No. 3.1e-58;
Matches 131; Conservative 17; Mismatches 38; Indels 1; Gaps 1;

QY 11 STLAVCVDEVEFTNOEVEKEFGLFRTYDECVTFOLFKSFRVRINFSHPKSAARARIE 70

Db 47 TSLFACVHAEVFEAREQKRFREALFTIYDDVTFOLFKSFRVRINFSHPKSAARARIE 106

QY 71 LHETORFGKKLKYFAOVQTPETDGDKLHAPQAKOFLISPPSPVGMKPISDATPV 130

Db 107 LHETDFNGQKLKYFAOVQMSGEVRDXYLLPQPVKQFLISPPASPPVGMKQSEDMAPV 166

QY 131 LNYDLVAVAKLGPGEKYEELHAGTSTPSVVHVCDSDMEEDDPKTSPPKXIQTTRPG 190

Db 167 INVDLCANVSKLGPGEKYEELHAGTSTPSVVHVCDSEETEEBETK-NPKOKIAQTTRPD 225

QY 191 LPPSVSN 197

Db 226 PPTAALN 232

RESULT 14

US-09-614-474-11
; Sequence 11, Application US/09614474
; GENERAL INFORMATION:
; APPLICANT: Loring, Jeanne F.
; APPLICANT: Lingley, Debora W.
; APPLICANT: Edwards, Carla M.
; APPLICANT: Sreeter, David G.
; TITLE OF INVENTION: DOWN SYNDROME CRITICAL REGION 1-LIKE PROTEINS
; FILE REFERENCE: PC-0013 US
; CURRENT APPLICATION NUMBER: US/09/614,474
; CURRENT FILING DATE: 2000-07-11
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PERL Program
; SEQ ID NO 11
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: 96017919
US-09-614-474-11

Query Match

Best Local Similarity 63.3%; Score 658.5; DB 20; Length 241;
Best Local Similarity 70.1%; Pred. No. 3.1e-58;
Matches 131; Conservative 17; Mismatches 38; Indels 1; Gaps 1;

QY 11 STLAVCVDEVEFTNOEVEKEFGLFRTYDECVTFOLFKSFRVRINFSHPKSAARARIE 70

Db 47 TSLFACVHAEVFEAREQKRFREALFTIYDDVTFOLFKSFRVRINFSHPKSAARARIE 106

QY 71 LHETORFGKKLKYFAOVQTPETDGDKLHAPQAKOFLISPPSPVGMKPISDATPV 130

Db 107 LHETDFNGQKLKYFAOVQMSGEVRDXYLLPQPVKQFLISPPASPPVGMKQSEDMAPV 166

QY 131 LNYDLVAVAKLGPGEKYEELHAGTSTPSVVHVCDSDMEEDDPKTSPPKXIQTTRPG 190

Db 167 INVDLCANVSKLGPGEKYEELHAGTSTPSVVHVCDSEETEEBETK-NPKOKIAQTTRPD 225

QY 191 LPPSVSN 197

Db 226 PPTAALN 232

RESULT 15

US-09-782-953-21
; Sequence 21, Application US/09782953
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674P21
; CURRENT APPLICATION NUMBER: US/09/782,953
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; PRIOR FILING DATE: 2000-07-07

APPLICANT: Edwards, Carla M.
APPLICANT: Streeter, David G.
TITLE OF INVENTION: DOWN SYNDROME CRITICAL REGION 1-LIKE PROTEINS
FILE REFERENCE: PC-0013 US
CURRENT APPLICATION NUMBER: US/09/614,474
CURRENT FILING DATE: 2000-07-11
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PERL Program
SEQ ID NO 10
LENGTH: 192
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No: g1435040
US-09-614-474-10

Query Match 94.1%; Score 980; DB 20; Length 192;
Best Local Similarity 96.4%; Pred. No. 3.2e-91;
Matches 185; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 6 MDCDVSTLVACVVDVEFTNOEVEKEFGLFRTYDECVTFLFKSFRVRINFSHPKSA 65
DB 1 MDCDVSTLVACVVDVEFTNOEVEKEFGLFRTYDECVTFLFKSFRVRINFSHPKSA 60
QY 66 RARIELHETOPRGKKLKYFAOVOTPETDGDKLHAPPOAKQFLISPPSSPPVGMKPI 125
DB 61 RARIELHETOPRGKKLKYFAOVOTPETDGDKLHAPPOAKQFLISPPSSPPVGMKPI 120
QY 126 DATPVNLVDLLYAVAKLGPGEKYLHAGTSTPSVVHVCDSDMEBEDPKTSKPKIIQ 185
DB 121 DATPVNLVDLLYAVAKLGPGEKYLHAGTSTPSVVHVCDSDIEBEDPKTSKPKIIQ 180
QY 186 TRRRGLPPSVSN 197
DB 181 TRRRGLPPSVSN 192

RESULT 6
US-09-782-953-18
Sequence 18, Application US/09782953
GENERAL INFORMATION:
APPLICANT: WILLIAMS, R. SANDERS
APPLICANT: ROTHERMEL, BEVERLY
TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
FILE REFERENCE: UTS-674PZ1
CURRENT APPLICATION NUMBER: US/09/782,953
CURRENT FILING DATE: 2001-02-13
PRIOR APPLICATION NUMBER: 60/216,601
PRIOR FILING DATE: 2000-07-07
NUMBER OF SEQ ID NOS: 27
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 18
LENGTH: 192
TYPE: PRT
ORGANISM: Homo sapiens
US-09-782-953-18

Query Match 94.1%; Score 980; DB 21; Length 192;
Best Local Similarity 96.4%; Pred. No. 3.2e-91;
Matches 185; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 6 MDCDVSTLVACVVDVEFTNOEVEKEFGLFRTYDECVTFLFKSFRVRINFSHPKSA 65
DB 1 MDCDVSTLVACVVDVEFTNOEVEKEFGLFRTYDECVTFLFKSFRVRINFSHPKSA 60
QY 66 RARIELHETOPRGKKLKYFAOVOTPETDGDKLHAPPOAKQFLISPPSSPPVGMKPI 125
DB 61 RARIELHETOPRGKKLKYFAOVOTPETDGDKLHAPPOAKQFLISPPSSPPVGMKPI 120
QY 126 DATPVNLVDLLYAVAKLGPGEKYLHAGTSTPSVVHVCDSDMEBEDPKTSKPKIIQ 185

DB 121 DATPVNLVDLLYAVAKLGPGEKYLHAGTSTPSVVHVCDSDIEBEDPKTSKPKIIQ 180
QY 186 TRRRGLPPSVSN 197
DB 181 TRRRGLPPSVSN 192

RESULT 7
US-09-782-953-19
Sequence 19, Application US/09782953
GENERAL INFORMATION:
APPLICANT: WILLIAMS, R. SANDERS
APPLICANT: ROTHERMEL, BEVERLY
TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
FILE REFERENCE: UTS-674PZ1
CURRENT APPLICATION NUMBER: US/09/782,953
CURRENT FILING DATE: 2001-02-13
PRIOR APPLICATION NUMBER: 60/216,601
PRIOR FILING DATE: 2000-07-07
NUMBER OF SEQ ID NOS: 27
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 19
LENGTH: 192
TYPE: PRT
ORGANISM: Homo sapiens
US-09-782-953-19

Query Match 94.1%; Score 980; DB 21; Length 192;
Best Local Similarity 96.4%; Pred. No. 3.2e-91;
Matches 185; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 6 MDCDVSTLVACVVDVEFTNOEVEKEFGLFRTYDECVTFLFKSFRVRINFSHPKSA 65
DB 1 MDCDVSTLVACVVDVEFTNOEVEKEFGLFRTYDECVTFLFKSFRVRINFSHPKSA 60
QY 66 RARIELHETOPRGKKLKYFAOVOTPETDGDKLHAPPOAKQFLISPPSSPPVGMKPI 125
DB 61 RARIELHETOPRGKKLKYFAOVOTPETDGDKLHAPPOAKQFLISPPSSPPVGMKPI 120
QY 126 DATPVNLVDLLYAVAKLGPGEKYLHAGTSTPSVVHVCDSDMEBEDPKTSKPKIIQ 185
DB 121 DATPVNLVDLLYAVAKLGPGEKYLHAGTSTPSVVHVCDSDIEBEDPKTSKPKIIQ 180
QY 186 TRRRGLPPSVSN 197
DB 181 TRRRGLPPSVSN 192

RESULT 8
US-10-030-613-1
Sequence 1, Application US/10030613
GENERAL INFORMATION:
APPLICANT: INCYTE GENOMICS, INC.
APPLICANT: TANG, Y. TOM
APPLICANT: YUE, HENRY
TITLE OF INVENTION: HUMAN PROTEINS INVOLVED IN DETOXIFICATION
FILE REFERENCE: PF-0711 PCT
CURRENT APPLICATION NUMBER: US/10/030,613
CURRENT FILING DATE: 2002-01-03
PRIOR APPLICATION NUMBER: 60/142,678
PRIOR FILING DATE: 1999-07-07
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PERL Program
SEQ ID NO 1
LENGTH: 234
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No: 3751586CD1
US-10-030-613-1

Qy	61	PKSAARARIELHETQFRGKKLYFAVQVPTETDQKLIHLAPQAPAKQFLISPPSSPPVG	120
Db	61	PKSAARARIELHETQFRGKKLYFAVQVPTETDQKLIHLAPQAPAKQFLISPPSSPPVG	120
Qy	121	WKPISDATPVNLVDLLYAVAKLGPGEKYEYELHAGTESTPSVVHVCDSDMEEEDPKTSPK	180
Db	121	WKPISDATPVNLVDLLYAVAKLGPGEKYEYELHAGTESTPSVVHVCDSDMEEEDPKTSPK	180
Qy	181	PKIIQTRRPGLPSPSVN	197
Db	181	PKIIQTRRPGLPSPSVN	197

```

RESULT 2
US-09-782-953-10
; Sequence 10, Application US/09782953
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; TITLE OF INVENTION: CALCIINEURIN INTERACTING PROTEIN (MCIP)
; FILE REFERENCE: UTSID:674P21
; CURRENT APPLICATION NUMBER: US/09782,953
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 607216,601
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 10
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-782-953-10

```

Query Match	100.0%;	Score 1041;	DB 21;	Length 197;
Best Local Similarity	100.0%;	Pred. No. 2e-97;		
Matches 197;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MPAPSMDCDVSTLVACVVDVVEFTNQEVKEAFEGFLRTYDCEVTFOQFKSPRRVRI	NSH	60
Db	1	MPAPSMDCDVSTLVACVVDVVEFTNQEVKEAFEGFLRTYDCEVTFOQFKSPRRVRI	NSH	60
Qy	61	PKSAARARIELHETQFPGKKLKYFAQVQTPETDGDKLHLAPPQAKQFLISPS	SSPPVG	120
Db	61	PKSAARARIELHETQFPGKKLKYFAQVQTPETDGDKLHLAPPQAKQFLISPS	SSPPVG	120
Qy	121	WKPIDATPVNLVDLLYAVAKLGPGEKYEHLHAGTSTPSVVVHVCDSDMEEEEDPK	TSK	180
Db	121	WKPIDATPVNLVDLLYAVAKLGPGEKYEHLHAGTSTPSVVVHVCDSDMEEEEDPK	TSK	180
Qy	181	PKIIOTRRRGCLPPSVSN		197
Db	181	PKIIOTRRRGCLPPSVSN		197

```

RESULT 3
PCT-US02-17382-132
; Sequence 132, Application PC/TUS0217382
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-062
; CURRENT APPLICATION NUMBER: PCT/US02/17382
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: US 60/296,076
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,605
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/357,253
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 234
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 132

```

```

; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-17382-132

```

```

RESULT 4
US-09-575-580B-6
; Sequence 6, Application US/09575580B
; GENERAL INFORMATION:
; APPLICANT: McKeon, P.
; APPLICANT: Kayako, K.
; APPLICANT: Ryeom, S.
; TITLE OF INVENTION: CALCIPRESSINS: ENDOGENOUS INHIBITORS OF CALCINEURIN
; TITLE OF INVENTION: USES AND REAGENTS RELATED THERETO
; FILE REFERENCE: HMV-048.01
; CURRENT APPLICATION NUMBER: US/09/575,580B
; CURRENT FILING DATE: 2000-05-22
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-575-580B-6

```

RESULT 5
US-09-614-474-10
; Sequence 10, Application US/09614474
; GENERAL INFORMATION:
; APPLICANT: Loring, Jeanne F.
; APPLICANT: Tingley, Debora W.

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Comugen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 11:38:44 ; Search time 135.874 Seconds
(without alignments)

934.784 Million cell updates/sec

Title: US-09-782-953-9

Sequence: 1 MPAPSMDCDVSTLVACVVDV.....SPKPKIQTRRPGLPPSVSN 197

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues 4569144

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database :

Pending_Patents_AA_Main:*

1: /cgn2_6/prodata/1/paa/PCTUS_COMB.pep:*

2: /cgn2_6/prodata/1/paa/US06_COMB.pep:*

3: /cgn2_6/prodata/1/paa/US07_COMB.pep:*

4: /cgn2_6/prodata/1/paa/US08_COMB.pep:*

5: /cgn2_6/prodata/1/paa/US081_COMB.pep:*

6: /cgn2_6/prodata/1/paa/US082_COMB.pep:*

7: /cgn2_6/prodata/1/paa/US083_COMB.pep:*

8: /cgn2_6/prodata/1/paa/US084_COMB.pep:*

9: /cgn2_6/prodata/1/paa/US085_COMB.pep:*

10: /cgn2_6/prodata/1/paa/US086_COMB.pep:*

11: /cgn2_6/prodata/1/paa/US087_COMB.pep:*

12: /cgn2_6/prodata/1/paa/US088_COMB.pep:*

13: /cgn2_6/prodata/1/paa/US089_COMB.pep:*

14: /cgn2_6/prodata/1/paa/US090_COMB.pep:*

15: /cgn2_6/prodata/1/paa/US091_COMB.pep:*

16: /cgn2_6/prodata/1/paa/US092_COMB.pep:*

17: /cgn2_6/prodata/1/paa/US093_COMB.pep:*

18: /cgn2_6/prodata/1/paa/US094_COMB.pep:*

19: /cgn2_6/prodata/1/paa/US095_COMB.pep:*

20: /cgn2_6/prodata/1/paa/US096_COMB.pep:*

21: /cgn2_6/prodata/1/paa/US097_COMB.pep:*

22: /cgn2_6/prodata/1/paa/US098_COMB.pep:*

23: /cgn2_6/prodata/1/paa/US099_COMB.pep:*

24: /cgn2_6/prodata/1/paa/US100_COMB.pep:*

25: /cgn2_6/prodata/1/paa/US101_COMB.pep:*

26: /cgn2_6/prodata/1/paa/US102_COMB.pep:*

27: /cgn2_6/prodata/1/paa/US60_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1041	100.0	197	21	US-09-782-953-9
2	1041	100.0	197	21	US-09-782-953-10
3	1020	98.0	197	1	PCT-US02-17382-132
4	993	95.4	192	19	US-09-575-5808-6
5	980	94.1	192	20	US-09-514-474-10
6	980	94.1	192	21	US-09-782-953-18

7	980	94.1	192	21	US-09-782-953-19	Sequence 19, Appl
8	905	86.9	234	24	US-10-030-613-1	Sequence 1, Appl1
9	905	86.9	234	27	US-60-142-678-1	Sequence 1, Appl1
10	905	86.9	234	25	US-10-104-047-2216	Sequence 2216, Ap
11	905	86.9	255	20	US-09-614-474-2	Sequence 2, Appl1
12	882	84.7	242	19	US-09-575-5808-5	Sequence 5, Appl1
13	658.5	63.3	241	1	PCT-US02-17382-133	Sequence 133, App
14	658.5	63.3	241	20	US-09-614-474-11	Sequence 11, Appl
15	658.5	63.3	241	21	US-09-782-953-21	Sequence 21, Appl
16	658.5	63.3	241	21	US-09-782-953-22	Sequence 22, Appl
17	646	62.1	236	19	US-09-575-5808-24	Sequence 24, Appl
18	641	61.6	197	19	US-09-575-5808-8	Sequence 8, Appl1
19	641	61.6	197	21	US-09-782-953-12	Sequence 12, Appl
20	641	61.6	197	21	US-09-782-953-13	Sequence 13, Appl
21	641	61.6	197	26	US-10-247-671-174	Sequence 174, App
22	641	61.6	197	27	US-09-782-953-24	Sequence 24, Appl
23	633.5	60.9	212	21	US-09-782-953-25	Sequence 25, Appl
24	633.5	60.9	212	21	US-09-782-953-5	Sequence 5, Appl1
25	632.5	60.8	198	21	US-09-782-953-7	Sequence 7, Appl1
26	632.5	60.8	198	21	US-10-104-047-2892	Sequence 2892, Ap
27	626.5	60.2	252	25	US-09-782-953-15	Sequence 15, Appl
28	622	59.8	197	21	US-09-782-953-16	Sequence 16, Appl
29	622	59.8	197	21	US-09-575-5808-4	Sequence 4, Appl1
30	613.5	58.9	198	19	US-09-782-953-3	Sequence 3, Appl1
31	613.5	58.9	198	21	US-09-782-953-4	Sequence 4, Appl1
32	613.5	58.9	198	21	US-09-782-953-7	Sequence 7, Appl1
33	596	57.3	170	19	US-09-575-5808-7	Sequence 7, Appl1
34	476.5	45.8	142	21	US-09-758-472-6466	Sequence 6466, Ap
35	476.5	45.8	142	26	US-10-235-926-6466	Sequence 6466, Ap
36	404	38.8	292	20	US-09-619-049-1440	Sequence 1440, Ap
37	404	38.8	292	20	US-60-167-374-118	Sequence 718, App
38	404	38.8	292	27	US-60-171-627-2109	Sequence 2109, App
39	404	38.8	292	27	US-60-173-386-180	Sequence 684, App
40	404	38.8	292	27	US-60-175-871-764	Sequence 764, App
41	404	38.8	292	27	US-60-184-775-692	Sequence 692, App
42	404	38.8	292	27	US-60-191-637-40818	Sequence 40818, A
43	404	38.8	292	27	US-60-191-637-40818	Sequence 744, App
44	404	38.8	292	27	US-60-191-700-744	Sequence 744, App
45	400.5	38.5	164	27	US-60-140-956-1718	Sequence 1718, Ap

ALIGNMENTS

RESULT 1

US-09-782-953-9 Application US/09782953

GENERAL INFORMATION:

APPLICANT: WILLIAMS, R. SANDERS

APPLICANT: ROTHERMEL, BEVERLY

TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE

TITLE OF INVENTION: CALCINEURIN INTERACTING PROTEIN (MCIP)

FILE REFERENCE: UTSD:674PZ1

CURRENT APPLICATION NUMBER: US/09/782,953

CURRENT FILING DATE: 2001-02-13

PRIOR APPLICATION NUMBER: 60/216,601

PRIOR FILING DATE: 2000-07-07

NUMBER OF SEQ ID NOS: 27

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 9

LENGTH: 197

TYPE: PRT

ORGANISM: Mus musculus

US-09-782-953-9

Query Match 100.0%; Score 1041; DB 21; Length 197;

Best Local Similarity 100.0%; Pred No. 26-97; 0; Indels 0; Gaps 0;

Matches 197; Conservative 0; Mismatches 0;

Db 1 MPAPSMDCDVSTLVACVVDVFTNQEVRKEKFGSLPFTYDECYTFQLFKSFRRVRINFSH 60

1 MPAPSMDCDVSTLVACVVDVFTNQEVRKEKFGSLPFTYDECYTFQLFKSFRRVRINFSH 60

THIS PAGE BLANK (USPTO)

Search completed: December 11, 2002, 11:50:15
Job time : 11.6273 secs

```
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 61358
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-61358
```

```
Query Match          61.6%; Score 641; DB 5; Length 197;
Best Local Similarity 64.2%; Pred. No. 2.6e-51;
Matches 120; Conservative 29; Mismatches 30; Indels 8; Gaps 2;

Qy 11 STLACVVDVEFTNQEVKEFGLFRITYDECVTQFOLFKSPRRVRINFSPKSAARARIE 70
Db 11 SSLIACVANSDFSESETRAKFESLFRITYDKDITFYFKSRVRINFSPNPFSAADARLQ 70

Qy 71 LHETOFRGKKLYFAQVOTPETDGDKLHLAPPQAPAKOFLISPPSPVGVKWPISDATPV 130
Db 71 LHKTEFLGKMKLYFAQTLHGSS---HLAPPNPDQFLISPPASPPVGWKQVEDATPV 126

Qy 131 LNYDLLYAVAKLGPGEKVELHAGTSTPSVVHVHVCDSMEEEEDPKTS-----PKPKIIQT 186
Db 127 INYDLLYAIKLGPGKEKVELHAATDTTSPVVHVHVCESDQKEEEMERMRPKPKIIQT 186

Qy 187 RRGGLPP 193
Db 187 RRPEYTP 193
```

RESULT 10

```
US-09-724-676A-61359
; Sequence 61359, Application US/09724676A
; GENERAL INFORMATION:
```

```
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724.676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 61359
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-61359
```

```
Query Match          61.6%; Score 641; DB 5; Length 197;
Best Local Similarity 64.2%; Pred. No. 2.6e-51;
Matches 120; Conservative 29; Mismatches 30; Indels 8; Gaps 2;

Qy 11 STLACVVDVEFTNQEVKEFGLFRITYDECVTQFOLFKSPRRVRINFSPKSAARARIE 70
Db 11 SSLIACVANSDFSESETRAKFESLFRITYDKDITFYFKSRVRINFSPNPFSAADARLQ 70

Qy 71 LHETOFRGKKLYFAQVOTPETDGDKLHLAPPQAPAKOFLISPPSPVGVKWPISDATPV 130
Db 71 LHKTEFLGKMKLYFAQTLHGSS---HLAPPNPDQFLISPPASPPVGWKQVEDATPV 126

Qy 131 LNYDLLYAVAKLGPGEKVELHAGTSTPSVVHVHVCDSMEEEEDPKTS-----PKPKIIQT 186
Db 127 INYDLLYAIKLGPGKEKVELHAATDTTSPVVHVHVCESDQKEEEMERMRPKPKIIQT 186

Qy 187 RRGGLPP 193
Db 187 RRPEYTP 193
```

RESULT 11

```
US-09-724-676A-61367
```

```
; Sequence 61367, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
```

```
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724.676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 61367
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-61367
```

```
Query Match          61.6%; Score 641; DB 5; Length 197;
Best Local Similarity 64.2%; Pred. No. 2.6e-51;
Matches 120; Conservative 29; Mismatches 30; Indels 8; Gaps 2;

Qy 11 STLACVVDVEFTNQEVKEFGLFRITYDECVTQFOLFKSPRRVRINFSPKSAARARIE 70
Db 11 SSLIACVANSDFSESETRAKFESLFRITYDKDITFYFKSRVRINFSPNPFSAADARLQ 70

Qy 71 LHETOFRGKKLYFAQVOTPETDGDKLHLAPPQAPAKOFLISPPSPVGVKWPISDATPV 130
Db 71 LHKTEFLGKMKLYFAQTLHGSS---HLAPPNPDQFLISPPASPPVGWKQVEDATPV 126

Qy 131 LNYDLLYAVAKLGPGEKVELHAGTSTPSVVHVHVCDSMEEEEDPKTS-----PKPKIIQT 186
Db 127 INYDLLYAIKLGPGKEKVELHAATDTTSPVVHVHVCESDQKEEEMERMRPKPKIIQT 186

Qy 187 RRGGLPP 193
Db 187 RRPEYTP 193
```

RESULT 12

```
US-09-724-676-61352
; Sequence 61352, Application US/09724676
; GENERAL INFORMATION:
```

```
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724.676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 61352
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-61352
```

```
Query Match          59.8%; Score 622; DB 5; Length 197;
Best Local Similarity 63.6%; Pred. No. 1.4e-49;
Matches 119; Conservative 26; Mismatches 34; Indels 8; Gaps 2;

Qy 11 STLACVVDVEFTNQEVKEFGLFRITYDECVTQFOLFKSPRRVRINFSPKSAARARIE 70
Db 11 SATTACHLDRPFDVGICRAKFSLFRTYDKDITFYFKSRVRINFSPNPFSAADARLQ 70

Qy 71 LHETOFRGKKLYFAQVOTPETDGDKLHLAPPQAPAKOFLISPPSPVGVKWPISDATPV 130
Db 71 LHKTEFLGKMKLYFAQTLHGSS---HLAPPNPDQFLISPPASPPVGWKQVEDATPV 126

Qy 131 LNYDLLYAVAKLGPGEKVELHAGTSTPSVVHVHVCDSMEEEEDPKTS-----PKPKIIQT 186
Db 127 INYDLLYAIKLGPGKEKVELHAATDTTSPVVHVHVCESDQKEEEMERMRPKPKIIQT 186

Qy 187 RRGGLPP 193
Db 187 RRPEYTP 193
```

RESULT 13

```
US-09-724-676A-61352
```

```
; Sequence 61352, Application US/09724676A
```



```

; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: Patent version 3.2
; SEQ ID NO 63940
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-63940

Query Match      88.7%; Score 923; DB 5; Length 184;
Best Local Similarity 95.6%; Pred. No. 4e-77;
Matches 175; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 6 MDCDVSTLVACVVDVEFTNQEVKEKFEGLFRTYDCCVTFQLFKSFRRVRINFSPKSA 65
DB 1 MDCDVSTLVACVVDVEFTNQEVKEKFEGLFRTYDCCVTFQLFKSFRRVRINFSPKSA 60

QY 66 RARIELHETQFRGKKLYFAQVQTPETDGDGKLHLAPPQAKQFLISPPSPVGVKPKIS 125
DB 61 RARIELHETQFRGKKLYFAQVQTPETDGDGKLHLAPPQAKQFLISPPSPVGVKPKIN 120

QY 126 DATPVNLYDLYAVAKLGPGEKVELHAGTSTPSVVHVCDSDMEEDDPKTSKPKEIIQ 185
DB 121 DATPVNLYDLYAVAKLGPGEKVELHAGTSTPSVVHVCDSDIEEDDPKTSKPKEIIQ 180

QY 186 TRR 188
DB 181 GDR 183

RESULT 3
US-09-724-676A-63940
; Sequence 63940, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: Patent version 3.2
; SEQ ID NO 63940
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-63940

Query Match      88.7%; Score 923; DB 5; Length 184;
Best Local Similarity 95.6%; Pred. No. 4e-77;
Matches 175; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 6 MDCDVSTLVACVVDVEFTNQEVKEKFEGLFRTYDCCVTFQLFKSFRRVRINFSPKSA 65
DB 1 MDCDVSTLVACVVDVEFTNQEVKEKFEGLFRTYDCCVTFQLFKSFRRVRINFSPKSA 60

QY 66 RARIELHETQFRGKKLYFAQVQTPETDGDGKLHLAPPQAKQFLISPPSPVGVKPKIS 125
DB 61 RARIELHETQFRGKKLYFAQVQTPETDGDGKLHLAPPQAKQFLISPPSPVGVKPKIN 120

QY 126 DATPVNLYDLYAVAKLGPGEKVELHAGTSTPSVVHVCDSDMEEDDPKTSKPKEIIQ 185
DB 121 DATPVNLYDLYAVAKLGPGEKVELHAGTSTPSVVHVCDSDIEEDDPKTSKPKEIIQ 180

QY 186 TRR 188
DB 181 GDR 183

US-09-724-676A-63940

Query Match      88.7%; Score 923; DB 5; Length 184;
Best Local Similarity 95.6%; Pred. No. 4e-77;
Matches 175; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 6 MDCDVSTLVACVVDVEFTNQEVKEKFEGLFRTYDCCVTFQLFKSFRRVRINFSPKSA 65
DB 1 MDCDVSTLVACVVDVEFTNQEVKEKFEGLFRTYDCCVTFQLFKSFRRVRINFSPKSA 60

QY 66 RARIELHETQFRGKKLYFAQVQTPETDGDGKLHLAPPQAKQFLISPPSPVGVKPKIS 125
DB 61 RARIELHETQFRGKKLYFAQVQTPETDGDGKLHLAPPQAKQFLISPPSPVGVKPKIN 120

QY 126 DATPVNLYDLYAVAKLGPGEKVELHAGTSTPSVVHVCDSDMEEDDPKTSKPKEIIQ 185
DB 121 DATPVNLYDLYAVAKLGPGEKVELHAGTSTPSVVHVCDSDIEEDDPKTSKPKEIIQ 180

QY 186 TRR 188
DB 181 GDR 183
```

```

RESULT 4
US-10-290-438-2
; Sequence 2, Application US/10290438
; GENERAL INFORMATION:
; APPLICANT: Loring, Jeanne F.
; APPLICANT: Tingley, Debora W.
; APPLICANT: Edwards, Carla M.
; APPLICANT: Streeter, David G.
; TITLE OF INVENTION: DOWN SYNDROME CRITICAL REGION 1-LIKE PROTEIN
; FILE REFERENCE: PC-0013-1CIP
; CURRENT APPLICATION NUMBER: US/10/290,438
; CURRENT FILING DATE: 2002-11-06
; PRIOR FILING DATE: 2000-07-11
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PERL Program
; SEQ ID NO 2
; LENGTH: 255
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 247500.5
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 11
; OTHER INFORMATION: unknown or other
US-10-290-438-2

Query Match      86.9%; Score 905; DB 6; Length 255;
Best Local Similarity 91.4%; Pred. No. 2.6e-75;
Matches 171; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

QY 11 STLVAQVVDVEFTNQEVKEKFEGLFRTYDCCVTFQLFKSFRRVRINFSPKSAARARIE 70
DB 69 NSLFACNVHQSVEGESEKFEGLFRTYDCCVTFQLFKSFRRVRINFSPKSAARARIE 128

QY 71 LHETQFRGKKLYFAQVQTPETDGDGKLHLAPPQAKQFLISPPSPVGVKPKISDATPV 130
DB 129 LHETQFRGKKLYFAQVQTPETDGDGKLHLAPPQAKQFLISPPSPVGVKPKINDATPV 188

QY 131 LNYDLLYAVAKLGPGEKVELHAGTSTPSVVHVCDSDMEEDDPKTSKPKEIIQTRRP 190
DB 189 LNYDLLYAVAKLGPGEKVELHAGTSTPSVVHVCDSDIEEDDPKTSKPKEIIQTRRP 248

QY 191 LPPSVSN 197
DB 249 LPPSVSN 255

RESULT 5
US-10-290-438-11
; Sequence 11, Application US/10290438
; GENERAL INFORMATION:
; APPLICANT: Loring, Jeanne F.
; APPLICANT: Tingley, Debora W.
; APPLICANT: Edwards, Carla M.
; APPLICANT: Streeter, David G.
; TITLE OF INVENTION: DOWN SYNDROME CRITICAL REGION 1-LIKE PROTEIN
; FILE REFERENCE: PC-0013-1CIP
; CURRENT APPLICATION NUMBER: US/10/290,438
; CURRENT FILING DATE: 2002-11-06
; PRIOR FILING DATE: 2000-07-11
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PERL Program
; SEQ ID NO 11
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 96017919
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 11:39:24 ; Search time 11.6273 Seconds
(without alignments)
118.077 Million cell updates/sec

Title: US-09-782-953-9

Perfect score: 1041
Sequence: 1 MFAPSMDCVSTLVACVVDV.....SPKPIQTRRGLPPSVSN 197

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 233758 seqs, 65991041 residues

Total number of hits satisfying chosen parameters: 233758

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_New.*
1: /cgn2_6/prodata/1/paa/PCT_NEW_COMB.pep.*
2: /cgn2_6/prodata/1/paa/US06_NEW_COMB.pep.*
3: /cgn2_6/prodata/1/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/prodata/1/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/prodata/1/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/prodata/1/paa/US10_NEW_COMB.pep.*
7: /cgn2_6/prodata/1/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	980	94.1	192	US-10-290-438-10	Sequence 10, Appl
2	923	88.7	184	US-09-724-676-613540	Sequence 613540, A
3	923	88.7	184	US-09-724-676-613540	Sequence 613540, A
4	905	86.9	255	US-10-290-438-2	Sequence 2, Appl
5	658.5	63.3	241	US-10-290-438-11	Sequence 11, Appl
6	641	61.6	197	US-09-724-676-61358	Sequence 61358, A
7	641	61.6	197	US-09-724-676-61359	Sequence 61359, A
8	641	61.6	197	US-09-724-676-61357	Sequence 61357, A
9	641	61.6	197	US-09-724-676-61358	Sequence 61358, A
10	641	61.6	197	US-09-724-676-61359	Sequence 61359, A
11	641	61.6	197	US-09-724-676-61357	Sequence 61357, A
12	622	59.8	197	US-09-724-676-61352	Sequence 61352, A
13	622	59.8	197	US-09-724-676-61352	Sequence 61352, A
14	598	57.4	200	US-09-724-676-61356	Sequence 61356, A
15	598	57.4	200	US-09-724-676-61356	Sequence 61356, A
16	595	57.2	171	US-09-724-676-61354	Sequence 61354, A
17	595	57.2	171	US-09-724-676-61354	Sequence 61354, A
18	482.5	46.3	205	US-09-724-676-61360	Sequence 61360, A
19	482.5	46.3	205	US-09-724-676-61360	Sequence 61360, A
20	473.5	45.5	162	US-09-724-676-61351	Sequence 61351, A
21	473.5	45.5	162	US-09-724-676-61351	Sequence 61351, A
22	470	45.1	154	US-09-724-676-61361	Sequence 61361, A
23	470	45.1	154	US-09-724-676-61361	Sequence 61361, A
24	451	43.3	154	US-09-724-676-61353	Sequence 61353, A
25	451	43.3	154	US-09-724-676-61353	Sequence 61353, A
26	427	41.0	157	US-09-724-676-61357	Sequence 61357, A

27	427	41.0	157	US-09-724-676-61357	Sequence 61357, A
28	424	40.7	156	US-09-724-676-61355	Sequence 61355, A
29	424	40.7	156	US-09-724-676-61355	Sequence 61355, A
30	323	31.0	111	US-09-513-999C-5849	Sequence 5849, Ap
31	300	28.8	96	US-09-724-676-61363	Sequence 61363, A
32	300	28.8	96	US-09-724-676-61363	Sequence 61363, A
33	300	28.8	96	US-09-724-676-61363	Sequence 61363, A
34	300	28.8	96	US-09-724-676-61363	Sequence 61363, A
35	219	21.0	58	US-10-203-138A-12463	Sequence 12463, A
36	174	16.7	56	US-10-203-138A-11146	Sequence 11146, A
37	141.5	13.6	104	US-09-724-676-61364	Sequence 61364, A
38	141.5	13.6	104	US-09-724-676-61364	Sequence 61364, A
39	132.5	12.7	61	US-09-724-676-61365	Sequence 61365, A
40	132.5	12.7	61	US-09-724-676-61365	Sequence 61365, A
41	89	8.5	722	US-09-724-676-90346	Sequence 90346, A
42	89	8.5	722	US-09-724-676-90346	Sequence 90346, A
43	89	8.5	724	US-09-724-676-90347	Sequence 90347, A
44	89	8.5	724	US-09-724-676-90347	Sequence 90347, A
45	89	8.5	807	US-09-724-676-90339	Sequence 90339, A

ALIGNMENTS

```

RESULT 1
US-10-290-438-10
; Sequence 10, Application US/10290438
; GENERAL INFORMATION:
; APPLICANT: Loring, Jeanne F.
; APPLICANT: Tingley, Debora W.
; APPLICANT: Edwards, Carla M.
; APPLICANT: Streeter, David G.
; TITLE OF INVENTION: DOWN SYNDROME CRITICAL REGION 1-LIKE PROTEIN
; FILE REFERENCE: PC-0013-1CIP
; CURRENT APPLICATION NUMBER: US/10/290,438
; PRIOR FILING DATE: 2002-11-06
; PRIOR APPLICATION NUMBER: 09/614,474
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PERL Program
; SEQ ID NO 10
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: g1435040
US-10-290-438-10

Query Match      94.1%; Score 980; DB 6; Length 192;
Best Local Similarity 96.4%; Pred. No. 2.6e-82;
Matches 185; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 6 MDDVSTLVACVVDVENVTOEYKEKEFGFLFRYDSCVTFQLEKSPRRVAINFSHPASAA 65
DB 1 MDDVSTLVACVVDVENVTOEYKEKEFGFLFRYDSCVTFQLEKSPRRVAINFSHPASAA 60

QY 66 RARIELHETOPRGRKLLKLYPAQVOTPETDQDKLHAPPOKAPFLISPPSPVGMKPIIS 125
DB 61 RARIELHETOPRGRKLLKLYPAQVOTPETDQDKLHAPPOKAPFLISPPSPVGMKPIIS 120

QY 126 DATPVNLYDLLVAVAKLGPGEKYLHAGTSTSPVVHVCDSDMEEDBDKTSPPKIIQ 185
DB 121 DATPVNLYDLLVAVAKLGPGEKYLHAGTSTSPVVHVCDSDIEEDBDKTSPPKIIQ 180

QY 186 TRRPGGLPPSVSN 197
DB 181 TRRPGGLPPSVSN 192

RESULT 2
US-09-724-676-613540
; Sequence 613540, Application US/09724676

```

THIS PAGE BLANK (USPTO)

COUNTRY: USA
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/853,831
FILING DATE: 09-MAY-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 2592
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 243 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-853-831-2

Query Match 7.6%; Score 79; DB 3; Length 243;
Best Local Similarity 25.8%; Pred. No. 0.49;
Matches 39; Conservative 14; Mismatches 46; Indels 52; Gaps 7;

QY 71 LHETQFRGKKLKLYFAOVQTPETDG-----DKLHLPQPAKQFLISPPSSPPVGW 121
DB 43 LHE-----LYDLDTAPEDEPNEBAVSQIFPDSVMLA-VQEGIDLTFPPA----- 86
QY 122 KPISDATPVLYDLVAVAKLGEGEKYELHAGTESTPSVVVHVCD-----SDME 171
DB 87 -PGSPPEPHLSRQ-----PEQPEQRALGVSMNLPVEVIDLTCHBAGFPSPDDED 136
QY 172 EEDPKTSPKPK-----IIQTRRPGLPSPV 195
DB 137 EEGVSEPEPEPEPEPARPTRRPMAPAI 167

RESULT 15
PCT-US93-09774-2
Sequence 2, Application PC/TUS9309774
GENERAL INFORMATION:
APPLICANT: LA JOLLA CANCER RESEARCH FOUNDATION
TITLE OF INVENTION: METHOD OF INHIBITING REPLICATION OF
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSER: CAMPBELL AND FLORES
STREET: 4370 LA JOLLA VILLAGE DRIVE, SUITE 700
CITY: SAN DIEGO
STATE: CALIFORNIA
COUNTRY: UNITED STATES
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/09774
FILING DATE: 12-OCT-1992
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: PERKINS, SUSAN M.
REGISTRATION NUMBER: 36,405
REFERENCE/DOCKET NUMBER: PP-LJ 9770
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-535-9001

TELEFAX: 619-535-8949
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 243 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US93-09774-2

Query Match 7.6%; Score 79; DB 5; Length 243;
Best Local Similarity 25.8%; Pred. No. 0.49;
Matches 39; Conservative 14; Mismatches 46; Indels 52; Gaps 7;

QY 71 LHETQFRGKKLKLYFAOVQTPETDG-----DKLHLPQPAKQFLISPPSSPPVGW 121
DB 43 LHE-----LYDLDTAPEDEPNEBAVSQIFPDSVMLA-VQEGIDLTFPPA----- 86
QY 122 KPISDATPVLYDLVAVAKLGEGEKYELHAGTESTPSVVVHVCD-----SDME 171
DB 87 -PGSPPEPHLSRQ-----PEQPEQRALGVSMNLPVEVIDLTCHBAGFPSPDDED 136
QY 172 EEDPKTSPKPK-----IIQTRRPGLPSPV 195
DB 137 EEGVSEPEPEPEPEPARPTRRPMAPAI 167

Search completed: December 11, 2002, 11:39:18
Job time: 14.2951 secs

```
Best Local Similarity 25.8%; Pred. No. 0.49;
Matches 39; Conservative 14; Mismatches 46; Indels 52; Gaps 7;

Qy 71 LHETQFRGKKLYFAQVOTPETDG-----DKLHAPPQPAKQFLISPPSPPVGW 121
Db 43 LHE-----LYDLDTAPEDPNEEAVSQIFPDSVMLA-VQEGIDLLTTPPA----- 86

Qy 122 KPISDATPVNLYLLYAVAKLPGKGYELHAGTSTPSVVHVCD-----SDMEE 171
Db 87 -PGSPPEPHLSRQ-----PEQPEQALGPVSMNLPVEVIDLTCHAGFPSPDDED 136

Qy 172 BEDPKTSPKPK-----IIOTRRPGLPPSV 195
Db 137 BEGPVSEPEPEPEPARTRRPKMAPAI 167

RESULT 12
US-08-301-316B-2
; Sequence 2, Application US/08301316B
; Patent No. 5776743
; GENERAL INFORMATION:
; APPLICANT: Frisch, Steven M.
; TITLE OF INVENTION: Method of Sensitizing Tumor Cells with Adenovirus E1A
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 06-SEP-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 1115
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 243 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-301-316B-2

Query Match 7.6%; Score 79; DB 1; Length 243;
Best Local Similarity 25.8%; Pred. No. 0.49;
Matches 39; Conservative 14; Mismatches 46; Indels 52; Gaps 7;

Qy 71 LHETQFRGKKLYFAQVOTPETDG-----DKLHAPPQPAKQFLISPPSPPVGW 121
Db 43 LHE-----LYDLDTAPEDPNEEAVSQIFPDSVMLA-VQEGIDLLTTPPA----- 86

Qy 122 KPISDATPVNLYLLYAVAKLPGKGYELHAGTSTPSVVHVCD-----SDMEE 171
Db 87 -PGSPPEPHLSRQ-----PEQPEQALGPVSMNLPVEVIDLTCHAGFPSPDDED 136

Qy 172 BEDPKTSPKPK-----IIOTRRPGLPPSV 195
Db 137 BEGPVSEPEPEPEPARTRRPKMAPAI 167

RESULT 13
```

```
US-08-473-399B-2
; Sequence 2, Application US/08473399B
; Patent No. 5866550
; GENERAL INFORMATION:
; APPLICANT: Frisch, Steven M.
; TITLE OF INVENTION: REVERSE-TRANSFORMATION OF CANCER CELLS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/960,112
; FILING DATE: 13-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 1697
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 243 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-473-399B-2

Query Match 7.6%; Score 79; DB 2; Length 243;
Best Local Similarity 25.8%; Pred. No. 0.49;
Matches 39; Conservative 14; Mismatches 46; Indels 52; Gaps 7;

Qy 71 LHETQFRGKKLYFAQVOTPETDG-----DKLHAPPQPAKQFLISPPSPPVGW 121
Db 43 LHE-----LYDLDTAPEDPNEEAVSQIFPDSVMLA-VQEGIDLLTTPPA----- 86

Qy 122 KPISDATPVNLYLLYAVAKLPGKGYELHAGTSTPSVVHVCD-----SDMEE 171
Db 87 -PGSPPEPHLSRQ-----PEQPEQALGPVSMNLPVEVIDLTCHAGFPSPDDED 136

Qy 172 BEDPKTSPKPK-----IIOTRRPGLPPSV 195
Db 137 BEGPVSEPEPEPEPARTRRPKMAPAI 167

RESULT 14
US-08-853-831-2
; Sequence 2, Application US/08853831
; Patent No. 6100243
; GENERAL INFORMATION:
; APPLICANT: Frisch, Steven M.
; TITLE OF INVENTION: Method of Sensitizing Tumor Cells
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
```

CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0333 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ. ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 559 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 496120
US-08-884-072-6

Query Match 7.8%; Score 81; DB 2; Length 559;
Best Local Similarity 25.4%; Pred. No. 0.97; Mismatches 63; Indels 46; Gaps 9;
Matches 44; Conservative 20;

QY 33 EQLFRTYD--ECVTFQFKSFRVRINFSHPKSARARIELHETQFRGKLLKYFAQVQT 90
DB 22 EGAFKASDQREMTPEHFQHLHEV--GYAAPSLPQTR-----RLRV----- 61
QY 91 PETDGDKHLAPQPAKQFLISPPSSPPVGWKPISDATPVLYND---LLYAVAKLGPG 146
DB 62 -DHSVTSLSH-DPLFEEQREVPQPPSP-----EDIPVEEDWPTFLPNVDKAGPAV 111
QY 147 KYELHAGTESTPVSVVVHVCSDMEEDPKTSPKPKTIO-----TRRPLGP 193
DB 112 PQBAIFLQKQPPQVHI-----EQKEIDPPAQPOEELVQKEVPHLAGQLPP 160

RESULT 10
US-09-212-168-6
Sequence 6, Application US/09212168
Patent No. 6303765
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: HUMAN EXTRACELLULAR MATRIX PROTEINS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: PaeSeQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/212,168
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/884,072
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0333 US

TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ. ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 559 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 496120
US-09-212-168-6

Query Match 7.8%; Score 81; DB 4; Length 559;
Best Local Similarity 25.4%; Pred. No. 0.97; Mismatches 63; Indels 46; Gaps 9;
Matches 44; Conservative 20;

QY 33 EQLFRTYD--ECVTFQFKSFRVRINFSHPKSARARIELHETQFRGKLLKYFAQVQT 90
DB 22 EGAFKASDQREMTPEHFQHLHEV--GYAAPSLPQTR-----RLRV----- 61
QY 91 PETDGDKHLAPQPAKQFLISPPSSPPVGWKPISDATPVLYND---LLYAVAKLGPG 146
DB 62 -DHSVTSLSH-DPLFEEQREVPQPPSP-----EDIPVEEDWPTFLPNVDKAGPAV 111
QY 147 KYELHAGTESTPVSVVVHVCSDMEEDPKTSPKPKTIO-----TRRPLGP 193
DB 112 PQBAIFLQKQPPQVHI-----EQKEIDPPAQPOEELVQKEVPHLAGQLPP 160

RESULT 11
US-07-960-112B-2
Sequence 2, Application US/07960112B
Patent No. 5516631
GENERAL INFORMATION:
APPLICANT: Frisch, Steven M.
TITLE OF INVENTION: Method of Inhibiting Replication of
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/960,112B
FILING DATE: 13-OCT-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 9429
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ. ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 243 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-960-112B-2

Query Match 7.6%; Score 79; DB 1; Length 243;

CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/574,959A
FILING DATE: 19-DEC-95
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: DFN-008
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 905 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-574-959A-9

Query Match
Best Local Similarity 29.6%; Score 89; DB 2; Length 905;
Matches 29; Conservative 13; Mismatches 44; Indels 12; Gaps 5;

QY 101 APPQPAKQFLISPPSPVGMKPISDATPVLYNDLYAVAKGPGKYEELHAGTESTPSV 160
DB 751 APPPTLPALPPESP-KVQPEPEPEP---GLLEVEE--PGTEEE--RADDTAFTL 801

QY 161 VHAVCDSDME---EEEDPKTSPKXIIQTRRPGLPSPV 195
DB 802 APEALPSQGEVEREGESPAGPPQELVEEPEXPFTL 839

RESULT 5
US-09-357-014-9
Sequence 9, Application US/09357014
Patent No. 6291645
GENERAL INFORMATION:
APPLICANT: Jaekyoon Shin, Inseil Young, Ratna K. Vadlamudi
TITLE OF INVENTION: P62 POLYPEPTIDES, RELATED POLYPEPTIDES
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/357,014
FILING DATE: 19-JUL-1999
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/574,959
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: DFN-008
TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 905 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 9:
US-09-357-014-9

Query Match
Best Local Similarity 29.6%; Score 89; DB 4; Length 905;
Matches 29; Conservative 13; Mismatches 44; Indels 12; Gaps 5;

QY 101 APPQPAKQFLISPPSPVGMKPISDATPVLYNDLYAVAKGPGKYEELHAGTESTPSV 160
DB 751 APPPTLPALPPESP-KVQPEPEPEP---GLLEVEE--PGTEEE--RADDTAFTL 801

QY 161 VHAVCDSDME---EEEDPKTSPKXIIQTRRPGLPSPV 195
DB 802 APEALPSQGEVEREGESPAGPPQELVEEPEXPFTL 839

RESULT 6
US-08-574-959A-7
Sequence 7, Application US/08574959A
Patent No. 5962224
GENERAL INFORMATION:
APPLICANT: Jaekyoon Shin, Inseil Young, Ratna K. Vadlamudi
TITLE OF INVENTION: P62 POLYPEPTIDES, RELATED POLYPEPTIDES
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/574,959A
FILING DATE: 19-DEC-95
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: DFN-008
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1135 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-574-959A-7

Query Match
Best Local Similarity 29.6%; Score 89; DB 2; Length 1135;
Matches 29; Conservative 13; Mismatches 44; Indels 12; Gaps 5;

QY 101 APPQPAKQFLISPPSPVGMKPISDATPVLYNDLYAVAKGPGKYEELHAGTESTPSV 160
DB 981 APPPTLPALPPESP-KVQPEPEPEP---GLLEVEE--PGTEEE--RADDTAFTL 1031

QY 161 VHAVCDSDME---EEEDPKTSPKXIIQTRRPGLPSPV 195

Query Match 56.3%; Score 586; DB 2; Length 171;
Best Local Similarity 66.5%; Pred. No. 1.7e-59;
Matches 113; Conservative 22; Mismatches 27; Indels 8; Gaps 2;

Qy 26 VKEKEGFRYDECVTFQFKSRFRVINFSPKSAARIELHETQFRGKLLKLYFAQ 87
Db 2 VYAKPESLFRYDRDITFYKSFKRVIRNFSPSAADARLQLHKTFLGKEMKLYFAQ 61
Qy 88 VQTPETDGDKLHAPPQAPKOFLLISPPSSPPVGMKPISDATPVNLVLLYAVAKLGPGKEK 147
Db 62 TLHIGSS---HLAPPNDKQFLISPPASPPVGYQVEDATPVNYDLYAISKLGPGEK 117
Qy 148 VELHAGTSTSPVVHVCDSDMEBEDPKTS-----PKKLIQTRRPGLP 193
Db 118 VELHAATDTTPSVVVHVCDSDQEEBEMERMRRPKKIIQTRRPEYTP 167

RESULT 2
US-08-348-518C-4
; Sequence 4, Application US/08348518C
; Patent No. 6022740
; GENERAL INFORMATION:
; APPLICANT: SUDOL, MARIUS
; APPLICANT: PEER, BORK
; APPLICANT: HENRY, CHEN
; TITLE OF INVENTION: A SH2 DOMAIN ASSOCIATED PROTEIN, A
; TITLE OF INVENTION: SIGNALLING DOMAIN THEREIN, NUCLEIC ACIDS ENCODING THE
; TITLE OF INVENTION: PROTEIN AND THE DOMAIN, AND DIAGNOSTIC AND THERAPEUTIC USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/348,518C
; FILING DATE: 01-DEC-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 454 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-348-518C-4

Query Match 9.7%; Score 101; DB 3; Length 454;
Best Local Similarity 32.1%; Pred. No. 0.0036;
Matches 34; Conservative 10; Mismatches 30; Indels 32; Gaps 5;

Qy 102 PPQPAKQFLISPPSSPPVGMKPISDATPVNLVLLYAVAKLGPGKEKVELHAGTSTSPV- 160
Db 8 PPQPAQOGQGGPPSQPPQGGPPS-----GPGQ--PAPAATQAAPQAP 48
Qy 161 -----VVHV-CDSDMEEEE-----DPKTSKPKLIQTRRPGLP 194
Db 49 PAGHQIVHVRGDSQETDLEALFNAMNPKTANVPQTPMRLKLPDS 94

US-08-348-518C-4

Query Match 9.7%; Score 101; DB 3; Length 454;
Best Local Similarity 32.1%; Pred. No. 0.0036;
Matches 34; Conservative 10; Mismatches 30; Indels 32; Gaps 5;

Qy 102 PPQPAKQFLISPPSSPPVGMKPISDATPVNLVLLYAVAKLGPGKEKVELHAGTSTSPV- 160
Db 8 PPQPAQOGQGGPPSQPPQGGPPS-----GPGQ--PAPAATQAAPQAP 48
Qy 161 -----VVHV-CDSDMEEEE-----DPKTSKPKLIQTRRPGLP 194
Db 49 PAGHQIVHVRGDSQETDLEALFNAMNPKTANVPQTPMRLKLPDS 94

RESULT 3
US-08-476-509B-4
; Sequence 4, Application US/08476509B
; Patent No. 6034212
; GENERAL INFORMATION:
; APPLICANT: SUDOL, MARIUS
; APPLICANT: PEER, BORK
; APPLICANT: HENRY, CHEN
; TITLE OF INVENTION: A SH2 DOMAIN ASSOCIATED PROTEIN, A
; TITLE OF INVENTION: SIGNALLING DOMAIN THEREIN, NUCLEIC ACIDS ENCODING THE
; TITLE OF INVENTION: PROTEIN AND THE DOMAIN, AND DIAGNOSTIC AND THERAPEUTIC USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,509B
; FILING DATE: 01-DEC-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-101 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 454 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-476-509B-4

Query Match 9.7%; Score 101; DB 3; Length 454;
Best Local Similarity 32.1%; Pred. No. 0.0036;
Matches 34; Conservative 10; Mismatches 30; Indels 32; Gaps 5;

Qy 102 PPQPAKQFLISPPSSPPVGMKPISDATPVNLVLLYAVAKLGPGKEKVELHAGTSTSPV- 160
Db 8 PPQPAQOGQGGPPSQPPQGGPPS-----GPGQ--PAPAATQAAPQAP 48
Qy 161 -----VVHV-CDSDMEEEE-----DPKTSKPKLIQTRRPGLP 194
Db 49 PAGHQIVHVRGDSQETDLEALFNAMNPKTANVPQTPMRLKLPDS 94

RESULT 4
US-08-574-959A-9
; Sequence 9, Application US/08574959A
; Patent No. 5962224
; GENERAL INFORMATION:
; APPLICANT: Jaekyoon Shin, Insil Joong, Ratna K. Vadlamudi
; APPLICANT: and Jack L. Strominger
; TITLE OF INVENTION: P62 POLYPEPTIDES, RELATED POLYPEPTIDES
; TITLE OF INVENTION: AND USES THEREFOR
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Comugen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 11:37:41 ; Search time 11.2951 Seconds
(without alignments)
513.170 Million cell updates/sec

Title: US-09-782-953-9

Perfect score: 1041

Sequence: 1 MPAPSMDCVSTLVACVVDV.....SPKXITITRRPGLPPSVSN 197

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*

1: /cgn2_6/prodata/1/aa/5A COMB.pep.*

2: /cgn2_6/prodata/1/aa/5B COMB.pep.*

3: /cgn2_6/prodata/1/aa/6A COMB.pep.*

4: /cgn2_6/prodata/1/aa/6B COMB.pep.*

5: /cgn2_6/prodata/1/aa/PCITUS COMB.pep.*

6: /cgn2_6/prodata/1/aa/backfillseq1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	586	56.3	171	2	US-08-665-040-2
2	101	9.7	454	3	US-08-348-518C-4
3	101	9.7	454	3	US-08-476-509B-4
4	89	8.5	905	2	US-08-574-959A-9
5	89	8.5	905	4	US-09-357-014-9
6	89	8.5	1135	2	US-08-574-959A-7
7	89	8.5	1135	4	US-09-357-014-9
8	81	7.8	434	2	US-08-815-718-3
9	81	7.8	559	2	US-08-884-072-6
10	81	7.8	559	4	US-09-212-168-6
11	79	7.6	243	1	US-07-560-112B-2
12	79	7.6	243	1	US-08-301-316B-2
13	79	7.6	243	2	US-08-473-399B-2
14	79	7.6	243	3	US-08-853-831-2
15	79	7.6	243	5	PCT-US93-09774-2
16	78	7.5	1018	1	US-08-408-093-6
17	78	7.5	1018	1	US-08-408-420A-6
18	78	7.5	1018	1	US-08-714-901-6
19	78	7.5	1018	1	US-08-040-741-6
20	76.5	7.3	330	4	US-09-145-391-2
21	76	7.3	154	4	US-07-945-283-2
22	75.5	7.3	154	4	US-08-866-928B-9
23	75	7.2	401	2	US-08-549-004A-5
24	75	7.2	401	2	US-09-051-982A-5
25	75	7.2	976	3	US-08-560-005-2
26	75	7.2	976	3	US-09-195-868-14
27	75	7.2	976	4	US-09-418-540-2

28	75	7.2	1187	4	US-08-664-962B-8	Sequence 8, App1
29	75	7.2	1187	4	US-09-311-743-8	Sequence 8, App1
30	75	7.2	1187	3	US-09-195-868-15	Sequence 15, App1
31	75	7.2	1229	3	US-09-195-868-28	Sequence 28, App1
32	75	7.2	2441	1	US-08-194-468-2	Sequence 2, App1
33	75	7.2	2441	3	US-08-961-739-2	Sequence 2, App1
34	75	7.2	2441	4	US-09-514-247A-8	Sequence 8, App1
35	72.5	7.0	659	4	US-09-562-737-12	Sequence 12, App1
36	71	6.8	350	2	US-08-960-023-8	Sequence 8, App1
37	71	6.8	357	1	US-08-145-006C-12	Sequence 12, App1
38	71	6.8	357	5	PCT-US94-00545-12	Sequence 12, App1
39	71	6.8	2972	4	US-09-579-181-2	Sequence 12, App1
40	71	6.8	3118	4	US-09-579-181-1	Sequence 1, App1
41	70.5	6.8	249	1	US-08-450-246-57	Sequence 57, App1
42	70.5	6.8	249	1	US-08-450-246-57	Sequence 57, App1
43	70.5	6.8	249	1	US-08-450-098-57	Sequence 57, App1
44	70.5	6.8	249	1	US-08-451-233-57	Sequence 57, App1
45	70.5	6.8	249	1	US-08-450-236-57	Sequence 57, App1

ALIGNMENTS

RESULT 1

US-08-665-040-2

Sequence 2, Application US/08665040

Patent No. 5869318

GENERAL INFORMATION:

APPLICANT: ESTIVILL, PALETA, XAVIER

APPLICANT: FUENTES, JUAN JOSE

APPLICANT: PRITCHARD, MELANIE

TITLE OF INVENTION: A NEW GENE SEQUENCE OF THE

TITLE OF INVENTION: DOWN SYNDROM CRITICAL REGION OF HUMAN CHROMOSOME 21,

TITLE OF INVENTION: IDENTIFIED BY A NEW "ALU-SPICING PCR" TECHNIQUE

TITLE OF INVENTION: CODING FOR A PROLINE-RICH PROTEIN (USC11) HIGHLY

TITLE OF INVENTION: EXPRESSED IN FETAL BRAIN AND IN HEART AND METHOD

TITLE OF INVENTION: FOR CHARACTERIZING IT.

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: LADAS & PARRY

STREET: 26 WEST 61ST STREET

CITY: NEW YORK

STATE: NEW YORK

COUNTRY: USA

ZIP: 10023

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK

COMPUTER: IBM PC COMPATIBLE

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WORD PERFECT 5.1 FOR DOS

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/665,040

FILING DATE: JUNE 7, 1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: ES P9501140

FILING DATE: JUNE 7, 1995

ATTORNEY/AGENT INFORMATION:

NAME: JANET I. CORD

REGISTRATION NUMBER: 33,778

REFERENCE/DOCKET NUMBER: U010815-9

TELEPHONE: (212) 708-1800

TELEFAX: (212) 246-8959

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 171 amino acids

TYPE: amino acid

STRANDEDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: peptide

US-08-665-040-2

```
Qy 110 ProAlaSerProProValGlyThrPheGlnValGluAspAlaThrProValIleAsnTyr 129
Db 335 CCCGCCCTCCGCCAGTGGGATGAACAGTGAAGATGCCGCCAGTCTATAACTAT 394
Qy 130 AspLeuLeuTyrAlaIleSerLysLeuGlyProGlyGluLysTyrGluLeuHisAlaAla 149
Db 395 GATCTTTATATGTCATCTCCAGCTGGGCGCAGGGGAAAGATGATGAATTCACGCCGCG 454
Qy 150 ThrAspProThrProSerValValHisValCysGluSerAspGlnGluAsnGluGlu 169
Db 455 ACTGACACCACTCCAGCGCTGGTGTCTCATGTATGTGAGAGTATCAAGAG--AAGGAG 511
Qy 170 GluGluGluGluMetGluArgMetLysArgProLysProLysPheGlnIleGlnThrArgArg 189
Db 512 GAAGAAGAGGAATGGAAGAATGAGGAGACCTAAGCCAAAATTTATCCAGCAGGAGG 571
Qy 190 ProGluTyrThrProIleHisLeuSer 198
Db 572 CCGAGTACACGCCGATCCACTCAGC 598

RESULT 14
US-09-724-676-12742
; Sequence 12742, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 12742
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676-12742

Alignment Scores:
Pred. No.: 1,04e-72 Length: 1056
Score: 731.00 Matches: 138
Percent Similarity: 97.24% Conservative: 3
Best Local Similarity: 95.17% Mismatches: 4
Query Match: 69.75% Indels: 0
DB: 5 Gaps: 0

US-09-782-953-3 (1-198) x US-09-724-676-12742 (1-1056)
Qy 1 MetGluGluValAspLeuGlnAspLeuProSerAlaThrIleAlaCysHisLeuAspPro 20
Db 25 ATGGAGGAGGTGGACCTCGAGGACCTGCCAGCGCCACCATCGCTGTACCTGACCCG 84
Qy 21 ArgValPheValAspGlyLeuCysArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40
Db 85 CGCGTGTTCGTGGAGCGCTGTGCCGGGCCAAATTTGAGTCCCTCTTTAGGACGTATGAC 144
Qy 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60
Db 145 AAGGACATCACCTTTTCAGTATTTTAAGAGCTTCAACAGAGTCAGATAAATTCAGCAAC 204
Qy 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
Db 205 CCCTTCTCCGACGACAGTCCAGGCTCCAGCTGCATAAGACTGAGTTTCTGGGAAGGAA 264
Qy 81 MetLysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProAsn 100
Db 265 ATGAAGTTATATTTTGTCTCAGACCTTACACATAGGAAGCTCACACCTGGCTCCGCCAAAT 324
Qy 101 ProAspLysGlnPheLeuIleSerProAlaSerProValGlyTrpLysGlnVal 120
Db 325 CCAGCAAGCAGAGTTTCTGATCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 384
Qy 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro 140
Db 385 GAAGATGCGACCCAGTCATTAATATGATCTCTTATATGCACTCCCAAGCTGGGGCCA 444
Qy 141 GlyGluLysTyrGlu 145
Db 445 GGTGATTGGCTAGAG 459

Search completed: December 15, 2002, 00:04:20
Job time : 57.7555 secs
```

```
Db 385 GAAGATGCGACCCAGTCATAAACTATGATCTCTTATATGCCATCTCAAGCTGGGGCCA 444
Qy 141 GlyGluLysTyrGlu 145
Db 445 GGTGATTGGCTAGAG 459

RESULT 15
US-09-724-676A-12742
; Sequence 12742, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 12742
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676A-12742

Alignment Scores:
Pred. No.: 1,04e-72 Length: 1056
Score: 731.00 Matches: 138
Percent Similarity: 97.24% Conservative: 3
Best Local Similarity: 95.17% Mismatches: 4
Query Match: 69.75% Indels: 0
DB: 5 Gaps: 0

US-09-782-953-3 (1-198) x US-09-724-676A-12742 (1-1056)
Qy 1 MetGluGluValAspLeuGlnAspLeuProSerAlaThrIleAlaCysHisLeuAspPro 20
Db 25 ATGGAGGAGGTGGACCTCGAGGACCTGCCAGCGCCACCATCGCTGTACCTGACCCG 84
Qy 21 ArgValPheValAspGlyLeuCysArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40
Db 85 CGCGTGTTCGTGGAGCGCTGTGCCGGGCCAAATTTGAGTCCCTCTTTAGGACGTATGAC 144
Qy 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60
Db 145 AAGGACATCACCTTTTCAGTATTTTAAGAGCTTCAACAGAGTCAGATAAATTCAGCAAC 204
Qy 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
Db 205 CCCTTCTCCGACGACAGTCCAGGCTCCAGCTGCATAAGACTGAGTTTCTGGGAAGGAA 264
Qy 81 MetLysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProAsn 100
Db 265 ATGAAGTTATATTTTGTCTCAGACCTTACACATAGGAAGCTCACACCTGGCTCCGCCAAAT 324
Qy 101 ProAspLysGlnPheLeuIleSerProAlaSerProValGlyTrpLysGlnVal 120
Db 325 CCAGCAAGCAGAGTTTCTGATCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 384
Qy 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro 140
Db 385 GAAGATGCGACCCAGTCATTAATATGATCTCTTATATGCACTCCCAAGCTGGGGCCA 444
Qy 141 GlyGluLysTyrGlu 145
Db 445 GGTGATTGGCTAGAG 459

Search completed: December 15, 2002, 00:04:20
Job time : 57.7555 secs
```



```
; SOFTWARE: PERL Program
; SEQ ID NO 60
; LENGTH: 2348
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 042176.5
US-10-240-965-60

Alignment Scores:
Pred. No.: 1.35e-86 Length: 2348
Score: 859.50 Matches: 168
Percent Similarity: 92.59% Conservative: 7
Best Local Similarity: 88.89% Mismatches: 12
Query Match: 82.01% Indels: 2
DB: 6 Gaps: 1

US-09-782-953-3 (1-198) x US-10-240-965-60 (1-2348)
Qy 11 SerAlaThrIleAlaCysHisLeuAspProArgValPheValAspGlyLeuCysArgAla 30
Db 181 AGCTCCCTGATTGCTGTGTGGCAACAGTGTATCTTTCAGGAAAGTGAAACCGAGGCC 240
Qy 31 LysPheGluSerLeuPheArgThrTyraPheLysAspThrPheGlnTyrrPheLysSer 50
Db 241 AAATTTGAGTCCCTCTTTAGGACGTATGACCAAGACATCACCCTTTTCAGTATTTTAAAGAGC 300
Qy 51 PheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAspAlaArgLeu 70
Db 301 TTCAAACGAGTCAGATAAATCTTTCAGCAACCCCTTCTCCGAGCAGATGCCAGGCTCCAG 360
Qy 71 LeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyrrPheAlaGlnThrLeuHis 90
Db 361 CTGCATAGACTGAGTTTCTGGGAAAGGAATGAAGTTATATTGCTCAGACCTTACAC 420
Qy 91 IleGlySerSerHisLeu-AlaProAsnProAspLysGlnPheLeuIleSerProPr 110
Db 421 ATAGGAAGCTCACACCTCGGCTCGGCCAAATCCAGCAAGCAGTTTCTGTATCTCCCTCC 480
Qy 110 oAlaSerProProValGlyTrrPheGlnValGluAspAlaThrProValIleAsnTyra 130
Db 481 CGCTCTCCGCGAGTGGATGGAACAAGTGAAGATCGACCCCGCATATAAATCTATGA 540
Qy 130 pLeuLeuTyrrAlaIleSerLysLeuGlyProGlyGluLysTyrrGluLeuHisAlaAlaTh 150
Db 541 TCTCTTATGCTCATCTCCAAAGCTGGGCGCCAGGGGAAAGATATGAATTTGCACGCGAGC 600
Qy 150 rAspProThrProSerValValHisValCysGluSerAspGlnGluAsnGluGlu 170
Db 601 TGACACCACTCCACGCGTGGTGGTCCATGTATGTGAGAGTGATCAAGAG---AAGGAGGA 657
Qy 170 uGluGluGluMetGluArgMetLysArgProLysProLysIleIleGlnThrArgAtqPr 190
Db 658 AGAAGAGGAATGGAAGAATGAGGAGACCTTAGCCAAAATTTATCCAGACGAGGAGGCC 717
Qy 190 oGluTyrrThrProIleHisLeuSer 198
Db 718 GGAGTACACGCGGATCCACCTCAGC 742

RESULT 10
US-09-724-676-12743
; Sequence 12743, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12743
; LENGTH: 1839
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676A-12743

Alignment Scores:
Pred. No.: 2.13e-85 Length: 1839
Score: 847.50 Matches: 162
Percent Similarity: 97.63% Conservative: 3
Best Local Similarity: 95.86% Mismatches: 3
Query Match: 80.87% Indels: 1
DB: 5 Gaps: 1

US-09-782-953-3 (1-198) x US-09-724-676-12743 (1-1839)
Qy 30 AlaLysPheGluSerLeuPheArgThrTyraPheLysAspThrPheGlnTyrrPheLys 49
Db 58 GCCAAATTTGAGTCCCTCTTTAGGACGTATGACAAAGACATCACCCTTTTCAGTATTTTAAAG 117
Qy 50 SerPheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAspAlaArgLeu 69
Db 118 AGCTTCAAACGAGTCAGATAAATCTTTCAGCAACCCCTTCTCCGAGCAGATGCCAGGCTC 177
Qy 70 ArgLeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyrrPheAlaGlnThrLeu 89
Db 178 CAGCTGCTAAGACTGAGTTTCTGGGAAAGGAATGAAGTTATATTGCTCAGACCTTA 237
Qy 90 HisIleGlySerSerHisLeuAlaProAsnProAspLysGlnPheLeuIleSerPro 109
Db 238 CACATAGAGAGCTCACACCTGGCTCCGCCAAATCCAGCAAGCAGTTTCTGTATCTCCCT 297
Qy 110 ProLysSerProProValGlyTrrPheGlnValGluAspAlaThrProValIleAsnTyrr 129
Db 298 CCCGCTCTCCGCGAGTGGATGGAACAAGTGAAGATCGACCCCGATCAATAACTAT 357
Qy 130 AspLeuLeuTyrrAlaIleSerLysLeuGlyProGlyGluLysTyrrGluLeuHisAlaAla 149
Db 358 GATCTCTTATATGCTCATCTCCAAAGCTGGGCGCCAGGGGAAAGATATGAATTTGCACGCG 417
Qy 150 ThrAspProThrProSerValValHisValCysGluSerAspGlnGluAsnGluGlu 169
Db 418 ACTGACACCACTCCACGCGTGGTGGTCCATGTATGTGAGAGTGATCAAGAG---AAGGAG 474
Qy 170 GluGluGluGluMetGluArgMetLysArgProLysProLysIleIleGlnThrArgArg 189
Db 475 GAAGAAGAGGAATGGAAGAATGAGGAGACCTAAGCAAAAATTTATCCAGACGAGGAGG 534
Qy 190 ProGluTyrrThrProIleHisLeuSer 198
Db 535 CCGGAGTACACGCGGATCCACCTCAGC 561

RESULT 11
US-09-724-676A-12743
; Sequence 12743, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12743
; LENGTH: 1839
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676A-12743

Alignment Scores:
Pred. No.: 2.13e-85 Length: 1839
Score: 847.50 Matches: 162
Percent Similarity: 97.63% Conservative: 3
Best Local Similarity: 95.86% Mismatches: 3
Query Match: 80.87% Indels: 1
DB: 5 Gaps: 1
```


Qy 11 SerAlaThrIleAlaCysHisLeuAspProArgValPheValAspGlyLeuCysArgAla 30
 Db 212 AGCTCCCTGATTGCTGTGGCAACAGTAGTATCTTTCAGCAAGAGTGAACACGAGGCC 271
 Qy 31 LysPheGluSerLeuPheArgThrTyraAspLysAspThrPheGlnTyraPheLysSer 50
 Db 272 AAATTTGAGTCCCTCTTTAGGACGATGACCAAGAGACATCACCCTTCAGTATTTTAAGAGC 331
 Qy 51 PheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAspAlaArgLeuArg 70
 Db 332 TTCAACAGAGTCAGATTAACCTTCAGCAACCCCTTCTCCGACGAGATGCCAGGCTCCAG 391
 Qy 71 LeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyraPheAlaGlnThrLeuHis 90
 Db 392 CTGCATAAGACTGAGTTTCTGGGAAAGGAATGAAGTTATATTTTGTCTCAGACCTTACAC 451
 Qy 91 IleGlySerSerHisLeuAlaProProAsnProAsnProAspLysGlnPheLeuLysSerProPro 110
 Db 452 ATAGGAAGCTCACCTCGCTCCGCAATCCAGCAAGCAGTTTCTGATCTCCCTCC 511
 Qy 111 AlaSerProProValGlyTrpLysGlnValGluAspAlaThrProValIleAsnTyraAsp 130
 Db 512 GCCTCTCCGCCAGTCGGATGGAACAGTGAAGATGCGACCCAGTCATTAACCTATGAT 571
 Qy 131 LeuLeuTyraAlaIleSerLysLeuGlyProGlyGluLysTyraGluLeuHisAlaAlaThr 150
 Db 572 CTCCTATATGCTCATCTCAAGCTGGGCGCAGGGGAAAAGTATGAATTCACGACGACT 631
 Qy 151 AspProThrProSerValValHisValCysGluSerAspGlnGluAsnGluGluGlu 170
 Db 632 GACACACTCCACGCGTGGTGTCTGATGTGAGAGTGAAGAGTGAAGAGGAA 688
 Qy 171 GluGluGluMetGluArgMetLysArgProLysProLysIleIleGlnThrArgArgPro 190
 Db 689 GAAGAGGAATGGAAGATGAGGAGACTTAAGCCAAAATTTATCCAGACGAGGAGCGC 748
 Qy 191 GluTyThrProIleHisLeuSer 198
 Db 749 GAGTACACGCCGATCCACCTCAGC 772

RESULT 6

US-09-724-676A-12756
 ; Sequence 12756, Application US/09724676A

; GENERAL INFORMATION:
 ; APPLICANT: Compugen LTD
 ; TITLE OF INVENTION: Variants of alternative splicing
 ; FILE REFERENCE: 129181.4 Compugen
 ; CURRENT APPLICATION NUMBER: US/09/724,676A
 ; CURRENT FILING DATE: 2000-11-28
 ; NUMBER OF SEQ ID NOS: 97222
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 12756
 ; LENGTH: 2050
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (7)-(7)
 ; OTHER INFORMATION: n is a,c,g, or t
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (47)-(47)
 ; OTHER INFORMATION: n is a,c,g, or t
 ; US-09-724-676A-12756

Alignment Scores:

Pred. No.: 3,97e-88 Length: 2050
 Score: 872.50 Matches: 168
 Percent Similarity: 93.09% Conservative: 7
 Best Local Similarity: 89.36% Mismatches: 12
 Query Match: 83.25% Indels: 1
 DB: 5 Gaps: 1

US-09-782-953-3 (1-198) x US-09-724-676A-12756 (1-2050)
 Qy 11 SerAlaThrIleAlaCysHisLeuAspProArgValPheValAspGlyLeuCysArgAla 30
 Db 212 AGCTCCCTGATTGCTGTGGCAACAGTAGTATCTTTCAGCAAGAGTGAACACGAGGCC 271
 Qy 31 LysPheGluSerLeuPheArgThrTyraAspLysAspThrPheGlnTyraPheLysSer 50
 Db 272 AAATTTGAGTCCCTCTTTAGGACGATGACCAAGAGACATCACCCTTCAGTATTTTAAGAGC 331
 Qy 51 PheLysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAspAlaArgLeuArg 70
 Db 332 TTCAACAGAGTCAGATTAACCTTCAGCAACCCCTTCTCCGACGAGATGCCAGGCTCCAG 391
 Qy 71 LeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyraPheAlaGlnThrLeuHis 90
 Db 392 CTGCATAAGACTGAGTTTCTGGGAAAGGAATGAAGTTATATTTTGTCTCAGACCTTACAC 451
 Qy 91 IleGlySerSerHisLeuAlaProProAsnProAspLysGlnPheLeuLysSerProPro 110
 Db 452 ATAGGAAGCTCACCTCGCTCCGCAATCCAGCAAGCAGTTTCTGATCTCCCTCC 511
 Qy 111 AlaSerProProValGlyTrpLysGlnValGluAspAlaThrProValIleAsnTyraAsp 130
 Db 512 GCCTCTCCGCCAGTCGGATGGAACAGTGAAGATGCGACCCAGTCATTAACCTATGAT 571
 Qy 131 LeuLeuTyraAlaIleSerLysLeuGlyProGlyGluLysTyraGluLeuHisAlaAlaThr 150
 Db 572 CTCCTATATGCTCATCTCAAGCTGGGCGCAGGGGAAAAGTATGAATTCACGACGACT 631
 Qy 151 AspProThrProSerValValHisValCysGluSerAspGlnGluAsnGluGluGlu 170
 Db 632 GACACACTCCACGCGTGGTGTCTGATGTGAGAGTGAAGAGTGAAGAGGAA 688
 Qy 171 GluGluGluMetGluArgMetLysArgProLysProLysIleIleGlnThrArgArgPro 190
 Db 689 GAAGAGGAATGGAAGATGAGGAGACTTAAGCCAAAATTTATCCAGACGAGGAGCGC 748
 Qy 191 GluTyThrProIleHisLeuSer 198
 Db 749 GAGTACACGCCGATCCACCTCAGC 772

RESULT 7

US-09-724-676-12747

; Sequence 12747, Application US/09724676
 ; GENERAL INFORMATION:
 ; APPLICANT: Compugen LTD
 ; TITLE OF INVENTION: Variants of alternative splicing
 ; FILE REFERENCE: 129181.4 Compugen
 ; CURRENT APPLICATION NUMBER: US/09/724,676
 ; CURRENT FILING DATE: 2000-11-28
 ; NUMBER OF SEQ ID NOS: 97222
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 12747
 ; LENGTH: 2297
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (11)-(11)
 ; OTHER INFORMATION: n is a,c,g, or t
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (18)-(18)
 ; OTHER INFORMATION: n is a,c,g, or t
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (23)-(23)
 ; OTHER INFORMATION: n is a,c,g, or t
 ; US-09-724-676-12747

Alignment Scores:

Pred. No.: 4.63e-88 Length: 2297

QY 51 PheLYsArGVaIARgIIeaSnPheSeSerAnProlSeSerAlaIAspAlaArgLeuArg 70
Db 213 TTCAAAGCGATGAGATAACTTCAGCAACCCCTTCTCCGACGAGATGCCAGGCTCCAG 272
QY 71 LeuHISLYsThrgIuPheLeuGlyLySGIuMeLYsLeuTYrPheAlaGlnThrLeuHIS 90
Db 273 CTGCATTAAGACTGAGATTTCTGGGAAAGAAATGAAGTTATTTTCTCCAGACCTTTACAC 332
QY 91 IIGLYSeSerHISLeuAlaProlProAnProlApLySGInPheLeuISeSerProPro 110
Db 333 ATAGAGAACTCACACTGCTCCGCCCAATCCAGACAGAGCTTCTGATCTCCCTCC 392
QY 111 AlAsErProProValGlyTYrLySGInValGluAspAlaThrProValIIeaSnTYrAsp 130
Db 393 GCCTTCGCCAGTGGGATGGAAACAGTGAAGATGGAGCCCGCATTAAGCTATGAT 452
QY 131 LeuLeuTYrAlaIleSeSerLYsLeuGlyProGlyGlyLYsTYrGlyLeuHISAlaIaThr 150
Db 453 CTCTTATATGCCATCTCCAGCTGGGCGGCAAGGAAAGTATGATTCACGCCAGCACT 512
QY 151 AsPProThrProSeSerValValHISValCYsGlySeSerAspGlnGluAsnGluGlu 170
Db 513 GACACCACTCCAGCGTGTGTCTCATGTATGTAGAGTCAAGAG--AAGGAGGAA 569
QY 171 GluGluGluMeGlyuArgMeLYsArgProLYsProLYsIleIleGlnThrArgArgPro 190
Db 570 GAAGAGGAAATGGAAAGATAGAGACTTAAGCCAAATATTCACAGACGAGAGCGCG 629
QY 191 GluTYrThrProIleHISLeuSer 198
Db 630 GAGTACAGCGCCGATCCACTCAGC 653

RESULT 4
US-09-724-676A-12748
; Sequence 12748, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12748
; LENGTH: 1931
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (11)..(11)
; OTHER INFORMATION: n is a,c,g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (18)..(18)
; OTHER INFORMATION: n is a,c,g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (23)..(23)
; OTHER INFORMATION: n is a,c,g, or t
US-09-724-676A-12748

Alignment Scores:
Pred. No.: 3,66e-88 Length: 1931
Score: 872.50 Matches: 168
Percent Similarity: 93.09% Conservative: 7
Best Local Similarity: 89.36% Mismatches: 12
Query Match: 83.25% Indels: 1
DB: 5 Gaps: 1

US-09-782-953-3 (1-198) x US-09-724-676A-12748 (1-1931)
QY 11 SeRAlArThrIIeaIaCYbHISLeuAsPProArGVaIPheValaAspGlyLeuCYsArgAla 30
||||: ||||||| :||| :|||

Db 93 AGCTCCCTGATTCCTGCTGTGGCAAAAGTGAATATCTTCAGCGAAAGTGAACAGGCGC 152
QY 31 LySPheGlySeSerLeuPheArgThrTYrAspLYsAspThrThrPheGlnTYrPheLYsSer 50
Db 153 AAATTTAGTCCCTCTTTAGAGCTATGACAGAGACATCCCTTTGAGTATTTTAAGAGC 212
QY 51 PheLYsArGVaIARgIIeaSnPheSeSerAnProlSeSerAlaIAspAlaArgLeuArg 70
Db 213 TTCAAAGCGATGAGATAACTTCAGCAACCCCTTCTCCGACGAGATGCCAGGCTCCAG 272
QY 71 LeuHISLYsThrgIuPheLeuGlyLySGIuMeLYsLeuTYrPheAlaGlnThrLeuHIS 90
Db 273 CTGCATTAAGACTGAGATTTCTGGGAAAGAAATGAAGTTATTTTCTCCAGACCTTTACAC 332
QY 91 IIGLYSeSerHISLeuAlaProlProAnProlApLySGInPheLeuISeSerProPro 110
Db 333 ATAGAGAACTCACACTGCTCCGCCCAATCCAGACAGAGCTTCTGATCTCCCTCC 392
QY 111 AlAsErProProValGlyTYrLySGInValGluAspAlaThrProValIIeaSnTYrAsp 130
Db 393 GCCTTCGCCAGTGGGATGGAAACAGTGAAGATGGAGCCCGCATTAAGCTATGAT 452
QY 131 LeuLeuTYrAlaIleSeSerLYsLeuGlyProGlyGlyLYsTYrGlyLeuHISAlaIaThr 150
Db 453 CTCTTATATGCCATCTCCAGCTGGGCGGCAAGGAAAGTATGATTCACAGCGAGCACT 512
QY 151 AsPProThrProSeSerValValHISValCYsGlySeSerAspGlnGluAsnGluGlu 170
Db 513 GACACCACTCCAGCGTGTGTCTCATGTATGTAGAGTCAAGAG--AAGGAGGAA 569
QY 171 GluGluGluMeGlyuArgMeLYsArgProLYsProLYsIleIleGlnThrArgArgPro 190
Db 570 GAAGAGGAAATGGAAAGATAGAGACTTAAGCCAAATATTCACAGACGAGAGCGCG 629
QY 191 GluTYrThrProIleHISLeuSer 198
Db 630 GAGTACAGCGCCGATCCACTCAGC 653

RESULT 5
US-09-724-676-12756
; Sequence 12756, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12756
; LENGTH: 2050
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (7)..(7)
; OTHER INFORMATION: n is a,c,g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (47)..(47)
; OTHER INFORMATION: n is a,c,g, or t
US-09-724-676-12756

Alignment Scores:
Pred. No.: 3,97e-88 Length: 2050
Score: 872.50 Matches: 168
Percent Similarity: 93.09% Conservative: 7
Best Local Similarity: 89.36% Mismatches: 12
Query Match: 83.25% Indels: 1
DB: 5 Gaps: 1

US-09-782-953-3 (1-198) x US-09-724-676-12756 (1-2050)


```
Qy 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValAlaGlnLeuAsnPheSerAsn 60
|||
Db 145 AAGGACATCACCTTTTCAAGAGCTTCAACAGAGTCAGAATAAACTTCAGCAAC 204
|||
Qy 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
|||
Db 205 CCCTTCTCCGAGCAGATGCCAGCTCCAGCTGCATAGACTGAGTTCTGGGAAGAA 264
|||
Qy 81 MetLysLeuTyrPheAlaGlnThrLeuHisLysGlySerSerHisLeuAlaProProAsn 100
|||
Db 265 ATGAAGTTATATTTTGTCTCAGACCTTACACATAGGAAGCTCACACCTGGCTCCGCCAAAT 324
|||
Qy 101 ProAspLysGlnPheLeuLysSerProProAlaSerProProValGlyTyrPheLysGlnVal 120
|||
Db 325 CCAGCAAGCAGATTTCTGATCTCCCTCCCGCTCTCCCGCAGTGGGATGGAACAAGTG 384
|||
Qy 121 GluAspAlaThrProValLeuAsnTyrAspLeuLeuTyrAlaLeuSerLysLeuGlyPro 140
|||
Db 385 GAAGATGCGACCCCGAGTCAAACTATGATCTCTTATATGCACTCCACAGCTGGGGCCA 444
|||
Qy 141 GlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160
|||
Db 445 GGGGAAAGTATGAATTCACGCGAGCTGACACCACTCCCGAGCTGGTGGTCCATGTA 504
|||
Qy 161 CysGluSerAspGlnGluAsnGluGluGluGluMetGluArgMetLysArgPro 180
|||
Db 505 TGTGAGAGTGTCAAGAG---AAGGAGGAAGAGGAAATGGAAAGAAATGAGGAGACCT 561
|||
Qy 181 LysProLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 198
|||
Db 562 AAGCCAAAATTTATCCAGACAGGAGCGGAGTAGTACACCGCATCCACCTCAGC 615
|||
RESULT 2
US-09-724-676A-12741
; Sequence 12741, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 12741
; LENGTH: 1893
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676A-12741
Alignment Scores:
Pred. No.: 2,298-102 Length: 1893
Score: 999.50 Matches: 190
Percent Similarity: 97.98% Conservative: 4
Best Local Similarity: 95.96% Mismatches: 3
Query Match: 95.37% Indels: 1
DB: 5 Gaps: 1
US-09-782-953-3 (1-198) x US-09-724-676A-12741 (1-1893)
Qy 1 MetGluGluValAspLeuGlnAspLeuProSerAlaThrIleAlaCysHisLeuAspPro 20
|||
Db 25 ATGGAGGAGGTGGACCTCGAGGACCTGCGCCAGCCACCATCGCTGTCCACCTGACCCG 84
|||
Qy 21 ArgValPheValAspGlyLeuCysArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40
|||
Db 85 CCGGTGTTCTGGACGGCTGTGCGCGGCAAAATTTAGTCCCTCTTTAGGACGATGAC 144
|||
Qy 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValAlaGlnLeuAsnPheSerAsn 60
|||
Db 145 AAGGACATCACCTTTTCAAGAGCTTCAACAGAGTCAGAATAAACTTCAGCAAC 204
|||
Qy 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
|||
```

```
Db 205 CCCTTCTCCGAGCAGATGCCAGCTCCAGCTGCATATAAGACTGAGTTTCTGGGAAAGGAA 264
|||
Qy 81 MetLysLeuTyrPheAlaGlnThrLeuHisLysGlySerSerHisLeuAlaProProAsn 100
|||
Db 265 ATGAAGTTATATTTTGTCTCAGACCTTACACATAGGAAGCTCACACCTGGCTCCGCCAAAT 324
|||
Qy 101 ProAspLysGlnPheLeuLysSerProProAlaSerProProValGlyTyrPheLysGlnVal 120
|||
Db 325 CCAGCAAGCAGATTTCTGATCTCCCTCCCGCTCTCCCGCAGTGGGATGGAACAAGTG 384
|||
Qy 121 GluAspAlaThrProValLeuAsnTyrAspLeuLeuTyrAlaLeuSerLysLeuGlyPro 140
|||
Db 385 GAAGATGCGACCCCGAGTCAAACTATGATCTCTTATATGCACTCCACAGCTGGGGCCA 444
|||
Qy 141 GlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160
|||
Db 445 GGGGAAAGTATGAATTCACGCGAGCTGACACCACTCCCGAGCTGGTGGTCCATGTA 504
|||
Qy 161 CysGluSerAspGlnGluAsnGluGluGluGluMetGluArgMetLysArgPro 180
|||
Db 505 TGTGAGAGTGTCAAGAG---AAGGAGGAAGAGGAAATGGAAAGAAATGAGGAGACCT 561
|||
Qy 181 LysProLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 198
|||
Db 562 AAGCCAAAATTTATCCAGACAGGAGCGGAGTAGTACACCGCATCCACCTCAGC 615
|||
RESULT 3
US-09-724-676-12748
; Sequence 12748, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 12748
; LENGTH: 1931
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676-12748
Alignment Scores:
Pred. No.: 3,668-88 Length: 1931
Score: 872.50 Matches: 168
Percent Similarity: 93.09% Conservative: 7
Best Local Similarity: 89.36% Mismatches: 12
Query Match: 83.25% Indels: 1
DB: 5 Gaps: 1
US-09-782-953-3 (1-198) x US-09-724-676-12748 (1-1931)
Qy 11 SerAlaThrIleAlaCysHisLeuAspProArgValPheValAspGlyLeuCysArgAla 30
|||
Db 93 AGCTCCCTCGATTTGCTGTGGCAACAGATGATCTTTCAGCGAAGTGAACACCGGCC 152
|||
Qy 31 LysPheGluSerLeuPheArgThrTyrAspLysAspThrThrPheGlnTyrPheLysSer 50
|||
Db 153 AAATTTGAGTCCCTCTTTAGGACGATATGACAGGACATCACCTTTTCAGTATTTAAGAGC 212
|||
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 14, 2002, 20:09:21 : Search time 52.7555 Seconds

(without alignments)
2732.361 Million cell updates/sec

Title: US-09-782-953-3

Sequence: 1 MEVDLDLPSATLCHLDP.....RPKKIKITRRPPYPIHLS 198

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 422006 seqs, 364007625 residues

Total number of hits satisfying chosen parameters: 844012

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=x1h
-Q=/cgn2_1/USPTO.spool/US09782953/runat_1122002_114431_17431/app_query.fasta_1.1173
-DB=Pending Patents NA New -QFMT=fastp -SUFFIX=rnpn -MINMATCH=0 -I -LOOPLC=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsun62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR.SCORB=pcp -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pcp -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USRR=US09782953 @CGN 1 1 70 @runat_1122002_114431_17431 -NCPUS=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -MAIT -LONGLOG -DEV TIMEOUT=120
-MAIN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELop=6 -DELext=7

Database : Pending_Patents_NA_New*

1: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:*
2: /cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq:*
3: /cgn2_6/ptodata/1/pna/US07_NEW_COMB.seq:*
4: /cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq:*
5: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:*
6: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:*
7: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	999.5	95.4	1893	US-09-724-676-12741	Sequence 12741, A
2	999.5	95.4	1893	US-09-724-676A-12741	Sequence 12741, A
3	872.5	83.3	1931	US-09-724-676-12748	Sequence 12748, A
4	872.5	83.3	1931	US-09-724-676A-12748	Sequence 12748, A
5	872.5	83.3	2050	US-09-724-676-12756	Sequence 12756, A
6	872.5	83.3	2050	US-09-724-676A-12756	Sequence 12756, A
7	872.5	83.3	2297	US-09-724-676-12747	Sequence 12747, A
8	872.5	83.3	2297	US-09-724-676A-12747	Sequence 12747, A
9	859.5	82.0	2348	US-10-240-965-60	Sequence 60, Appl
10	847.5	80.9	1839	US-09-724-676-12743	Sequence 12743, A
11	847.5	80.9	1839	US-09-724-676A-12743	Sequence 12743, A

12	847.5	80.9	1876	US-09-724-676-12745	Sequence 12745, A
13	847.5	80.9	1876	US-09-724-676A-12745	Sequence 12745, A
14	731	69.8	1056	US-09-724-676-12742	Sequence 12742, A
15	731	69.8	1056	US-09-724-676A-12742	Sequence 12742, A
16	620.5	57.2	3159	US-10-290-438-1	Sequence 1, Appl1
17	607	57.9	911	US-09-724-676-12749	Sequence 12749, A
18	607	57.9	911	US-09-724-676A-12749	Sequence 12749, A
19	604	57.6	1094	US-09-724-676-12750	Sequence 12750, A
20	604	57.6	1094	US-09-724-676A-12750	Sequence 12750, A
21	604	57.6	1213	US-09-724-676-12740	Sequence 12740, A
22	604	57.6	1213	US-09-724-676A-12740	Sequence 12740, A
23	579	55.2	1002	US-09-724-676-12744	Sequence 12744, A
24	579	55.2	1002	US-09-724-676A-12744	Sequence 12744, A
25	579	55.2	1039	US-09-724-676-12746	Sequence 12746, A
26	579	55.2	1039	US-09-724-676A-12746	Sequence 12746, A
27	578.5	55.2	2411	US-09-724-676-15329	Sequence 15329, A
28	578.5	55.2	2411	US-09-724-676A-15329	Sequence 15329, A
29	539	51.4	615	US-10-290-438-8	Sequence 8, Appl1
30	509.5	48.6	1021	US-10-290-438-6	Sequence 3, Appl1
31	439	41.9	442	US-09-513-999C-1772	Sequence 1772, Ap
32	415.5	39.6	1577	US-09-724-676-12755	Sequence 12755, A
33	415.5	39.6	1577	US-09-724-676A-12755	Sequence 12755, A
34	415.5	39.6	1943	US-09-724-676-12752	Sequence 12752, A
35	415.5	39.6	1943	US-09-724-676A-12752	Sequence 12752, A
36	285.5	27.2	446	US-10-203-138A-2101	Sequence 2101, Ap
37	266.5	25.4	486	US-10-203-138A-827	Sequence 827, Appl
38	198.5	18.9	531	US-10-290-438-6	Sequence 6, Appl1
39	187.5	17.9	123	US-09-620-607B-1320	Sequence 5980, Ap
40	156	14.9	465	US-09-724-676-12753	Sequence 1320, Ap
41	150	14.3	557	US-09-724-676-12753	Sequence 12753, A
42	150	14.3	557	US-09-724-676A-12753	Sequence 12753, A
43	147	14.0	740	US-09-724-676-12754	Sequence 12754, A
44	147	14.0	740	US-09-724-676A-12754	Sequence 12754, A
45	140	13.4	85	US-10-203-138A-7216	Sequence 7216, Ap

ALIGNMENTS

RESULT 1
US-09-724-676-12741
Sequence 12741, Application US/09724676
GENERAL INFORMATION:
APPLICANT: CompuGen LTD
TITLE OF INVENTION: Variants of alternative splicing
FILE REFERENCE: 129181.4 CompuGen
CURRENT APPLICATION NUMBER: US/09/724, 676
CURRENT FILING DATE: 2000-11-28
NUMBER OF SEQ ID NOS: 97222
SOFTWARE: PatentIn version 3.2
SEQ ID NO 12741
LENGTH: 1893
TYPE: DNA
ORGANISM: Homo sapiens
US-09-724-676-12741

Alignment Scores:

Pred. No.: 2,296-102
Score: 999.50
Percent Similarity: 97.98%
Best Local Similarity: 95.96%
Query Match: 95.37%
DB: 5
Length: 1893
Matches: 190
Conservative: 4
Mismatch: 3
Indels: 1
Gaps: 1

US-09-782-953-3 (1-198) x US-09-724-676-12741 (1-1893)

QY 1 MetGIGlivaIaApIeugInaApIeupProSerAlaThrIAlaCYeHIsIeuaAppro 20
DB 25 ATGAGAGAGTGGACCTCGAGGACCTGCCACGACCATCGCTGACCTGAGCCG 84
QY 21 ArgvalPhvaIaAsglIyuecyabaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 40
DB 85 CCGGtGTGTGGAGACCGCTGTGCGGCGCAATTGATTCCTCTTAAAGACGATATAC 144

THIS PAGE BLANK (series)

```
Db 364 CAGAAAAT---GGCTACGAGAAAAAGATTCCAGATGCAACCCCTCTCTTGATGCC 420
Qy 130 AspleuleuTyraIaIleSerLyLeuGIyProGIyGIuLySTyGIuLeuHleAlaIa 149
    |||
    ::|
    |||
Db 421 AGCTCACTGGAGTATCTGTTGCTCAGGA-----CAGTATGATTTGGTGAATGC 471
Qy 150 ThrAspProThrProSerValValIValHlValCyseGIuSerAspGIuGIuHnGIuGIu 169
    |||
    ::|
    |||
Db 472 CTGGCTCTCT-----ATTCGAGACCCCAAGACCGAGAGATGCACATGAT 516
Qy 170 GIuGIuGIuGIu-----MetGIu 175
    |||
    |||
Db 517 ATTCAGAACGAGTCTCTAGGAGTGGCTCTCTGSCCATCTCACACTATGCCATGATGAAG 576
Qy 176 ArgMetLyAlaArgProLyv---ProLyAlleIleGIuInThrArgArgProGIuTyThrPro 194
    ::|
    ::|
    ::|
Db 577 AAGATGCAGTTGCCAGAACTGCCCAAGACATCAGCTACAAAGCA-----TATATTCCA 630
```

Search completed: December 14, 2002, 22:30:09
Job time : 43.0506 secs

Db 919 AGGCATCCCTTCAAGCGCTGTGACAGATGACGTGCGGACCTGTGCAT 872
 RESULT 13
 US-08-097-997A-10
 Sequence 10, Application US/08097997A
 Patent No. 5728336
 GENERAL INFORMATION:
 APPLICANT: Ihle, James N.
 APPLICANT: Silvenoinen, Olli
 APPLICANT: Wiltuhn, Bruce A.
 APPLICANT: Queile, Frederick W.
 TITLE OF INVENTION: Jak Kinases and Regulation of Cytokine Signal
 TITLE OF INVENTION: Transduction
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Sterne, Kessler, Goldstein & Fox
 STREET: 1100 New York Avenue, Suite 600
 CITY: Washington
 STATE: D.C.
 COUNTRY: U.S.A.
 ZIP: 20005-3934
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/097,997A
 FILING DATE: 29-JULY-1994
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Fox, Samuel L.
 REGISTRATION NUMBER: 30,353
 REFERENCE/DOCKET NUMBER: 0656.0370000/SLF/GKT
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 371-2600
 TELEFAX: (202) 371-2540
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3429 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..3426
 US-08-097-997A-10
 Alignment Scores:
 Pred. No.: 10.6 Length: 3429
 Score: 79.00 Matches: 48
 Percent Similarity: 39.00% Conservative: 30
 Best Local Similarity: 24.00% Mismatches: 64
 Query Match: 7.54% Indels: 58
 DB: 1 Gaps: 13
 US-09-782-953-3 (1-198) x US-08-097-997A-10 (1-3429)
 QY 12 AATThTlEaJaCyehIsLeuaApProAenProApLyGlnPheLeuIleSerProAla 31
 Db 154 GCTGCACAGCAAGCGTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 186
 QY 32 PheGluSerLeuPheArgThrTyrAspLysAspThrThreGlnTyrPheLysSerPhe 51
 Db 187 ---CACAACTCTTTGGCCCTGTATGACGAGAACACCAAGCTCTGTGTGCTCCAAAT--- 240
 QY 52 LysArgValArgIleAsnPheSerAsnProLeuSerAlaAlaAspAlaArgLeuArgLeu 71
 Db 241 ---CGCACCATCAACCGTTGATGACAGAGATGCC-----CTCCGGCTC 279
 QY 72 HLeLysThrGlnPheLeuGlyLysGluMetLysLeuTyrPheAlaGlnThrLeuHisIle 91

Db 280 CAC-----TACCGATGAGTTCTATTTCAC----- 306
 QY 92 GlySerSerHisLeuAlaProAenProAenProApLyGlnPheLeuIleSerProAla 111
 Db 307 ---AATTGGCATGGAACCAACGCAATGAGCACTAGTGTGGGTATTCTCCAAAGAG 353
 QY 112 SerProProValGlyTrp-----LysGlnValGluAspAlaThrProValIleAsnTyr 129
 Db 364 CAGAAAAAT---GGCTACGAGAAAAAAGATTCCAGATGCAACCCCTCTCTGTATGCC 420
 QY 130 AspLeuLeuTyrAlaIleSerLysLeuGlyProGlyGlyLysTyrGlnLeuHisAlaIle 149
 Db 421 AGCTCACTGAGTATCTGTTTCTCAGGA-----CAGTATGATTTGTGTAATGC 471
 QY 150 ThrAspProThrProSerValValHisValCysGluSerAspGlnGluAsnGlu 169
 Db 472 CTGGCTCCT-----ATTCCAGACCCCAAGACGAGCAGATGACATGAT 516
 QY 170 GlnGlnGlnGlu-----MetGln 175
 Db 517 ATTGAGACGAGTGTAGGATGGCTGCTGGCCATCTCACACTATGCATGATGAG 576
 QY 176 ArgMetLysArgProLys---ProLysIleIleGlnThrArgArgProGluTyrThrPro 194
 Db 577 AGATGCAAGTTGCGCAAGCTGCCCAAGGACATGACTAGTCAAGCGA-----TATATTCCA 630
 RESULT 14
 US-08-665-574C-10
 Sequence 10, Application US/08665574C
 Patent No. 6136595
 GENERAL INFORMATION:
 APPLICANT: Ihle, James N.
 APPLICANT: Silvenoinen, Olli
 APPLICANT: Wiltuhn, Bruce A.
 TITLE OF INVENTION: Jak Kinases and Regulation of Cytokine
 TITLE OF INVENTION: Signal Transduction
 NUMBER OF SEQUENCES: 16
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Sterne, Kessler, Goldstein & Fox
 STREET: 1100 New York Avenue, Suite 600
 CITY: Washington
 STATE: D.C.
 COUNTRY: U.S.A.
 ZIP: 20005-3934
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/665,574C
 FILING DATE: 18-JUN-1996
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/282,012
 FILING DATE: 29-JUL-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/097,997
 FILING DATE: 29-JUL-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/118,968
 FILING DATE: 09-SEP-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Bugalsky, Lawrence B.
 REGISTRATION NUMBER: 35,086
 REFERENCE/DOCKET NUMBER: 0656.0370002/SLF/LBB
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 371-2600
 TELEFAX: (202) 371-2540
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3429 base pairs
 TYPE: nucleic acid

```
Db 1507 GATCATTTATCAGGATCATGCTTGTCTCTCTCTCTCAGAACTCATAGCA--- 1451
Qy 39 TyrAspLysAspThrPheGlnTyrPheLysSerPheLysArgValArgLysPhe 58
Db 1450 ---GAAACCCCTAGCAGCATGCAATACATCTCTCTCTGAAACTTATCTGAGAGGCTCT 1394
Qy 59 SerAsnProLeuSerAlaAlaAspAlaArgLeuLeu---HisLysThrGluPheLeu 77
Db 1393 GTTGAGTTATATACATCCAGTACGGCTTGAGTCGATGGCCCAAAACA--- 1343
Qy 78 GlyLysGluMetLysLeuTyrPhe--- 85
Db 1342 ---AAGAAGCTCTCCATCTTTCTCAAAGTTCTTAACACACTGGAGAGACATTGTATCC 1286
Qy 86 --- 86
Db 1285 ATGTAGCCTTAGTGTGCGAAAGCCATAGCTGTGTATCCATCCATCTTCACATTGGATT 1226
Qy 94 SerHisLeuAlaProProAsnProAspLysGlnPheLeuLeuSerPro--- 109
Db 1225 CCT---AGCCAGCAAAATCCC---ATCTTCAGTCCAGTGGCTGTTCAC 1184
Qy 110 --- 110
Db 1183 ATAGTCCACTGCTCTCAATCTCTGTTGGAAGTAGCGGAGATCCC--- 1136
Qy 119 GlnValGluAspAlaThrProValLeuLeuTyrAlaAlaSerLysLeu 138
Db 1135 ---GAGTCCTCAACCGATCCACACCCAGATGTCTCGAAAGATCCACCGGATA 1082
Qy 139 GlyProGluLys---TyrGluLeuHis 147
Db 1081 AACATTGGGGACTCCCGCTGAAATTTTGTGAAATCTCTGTGTGATGTATTCGAAGCACTT 1022
Qy 148 AlaAlaThrAspProThrProSerValValHisValCysGluSerAspGlnGluAsn 167
Db 1021 CTTGTACCGGT---TTGCATACGCGGTAAAGCGGTAGCGGAAGG 980
Qy 168 GluGluGluGluGluMetGluArgMetLysArgProLysProLysIleLeuGlnThr 187
Db 979 AGAGAACAGAGAGGAGCGCTGCAGAGCTGGAGGTTCAGAGCTCGGCCAGTCCAGGTC 920
Qy 188 ArgArgPro---GluTyrThrProLeuHis 196
Db 919 AGGCATCCCTCAAGCGCTGCAGGATGACGCTCGGGACCTGTGCAT 872
```

RESULT 12

US-08-405-254-7/c
Sequence 7, Application US/08405254
Patent No. 5773288

GENERAL INFORMATION:

APPLICANT: BRIGGS, STEVEN P.
APPLICANT: BENSON, ROBERT J.
TITLE OF INVENTION: PLANT GENES AFFECTING GIBBERELLIC ACID
NUMBER OF INVENTIONS: BIOSYNTHESIS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESS: FOLEY & LARDNER
STREET: 3000 K STREET, NW, SUITE 500
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20007
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/405,254
FILING DATE: 16-MAR-1995
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/261,465
FILING DATE: 17-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: BENT, STEPHEN A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 33229/299/PIHI
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2784 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-405-254-7

Alignment Scores:
Pred. No.: 7.75 Length: 2784
Score: 79.00 Matches: 49
Percent Similarity: 31.78% Conservative: 26
Best Local Similarity: 20.76% Mismatches: 79
Query Match: 7.54% Indels: 82
DB: 11 Gaps: 11

US-09-782-953-3 (1-198) x US-08-405-254-7 (1-2784)

```
Qy 19 AspProArgValPheValAspGlyLeuCysArgAlaLysPheGluSerLeuPheArgThr 38
Db 1507 GATCATTTATCAGGATCATGCTTGTCTCTCTCTCTGCTCAGAACTCATAGCA--- 1451
Qy 39 TyrAspLysAspThrPheGlnTyrPheLysSerPheLysArgValArgLysPhe 58
Db 1450 ---GAAACCCCTAGCAGCATGCAATACATCTCTCTCTGAAACTTATCTGAGAGGCTCT 1394
Qy 59 SerAsnProLeuSerAlaAlaAspAlaArgLeuLeu---HisLysThrGluPheLeu 77
Db 1393 GTTGAGTTATATACATCCAGTACGGCTTGAGTCGATGGCCCAAAACA--- 1343
Qy 78 GlyLysGluMetLysLeuTyrPhe--- 85
Db 1342 ---AAGAAGCTCTCCATCTTTCTCAAAGTTCTTAACACACTGGAGAGACATTGTATCC 1286
Qy 86 --- 86
Db 1285 ATGTAGCCTTAGTGTGCGAAAGCCATAGCTGTGTATCCATCTTCACATTGGATT 1226
Qy 94 SerHisLeuAlaProProAsnProAspLysGlnPheLeuLeuSerPro--- 109
Db 1225 CCT---AGCCAGCAAAATCCC---ATCTTCAGTCCAGTGGCTGTTCAC 1184
Qy 110 --- 110
Db 1183 ATAGTCCACTGCTCTCAATCTCTGTTGGAAGTAGCGGAGATCCC--- 1136
Qy 119 GlnValGluAspAlaThrProValLeuLeuTyrAlaAlaSerLysLeu 138
Db 1135 ---GAGTCCTCAACCGATCCACACCCAGATGTCTCGAAAGATCCACCGGATA 1082
Qy 139 GlyProGluLys---TyrGluLeuHis 147
Db 1081 AACATTGGGGACTCCCGCTGAAATTTTGTGAAATCTCTGTGTGATGTATTCGAAGCACTT 1022
Qy 148 AlaAlaThrAspProThrProSerValValHisValCysGluSerAspGlnGluAsn 167
Db 1021 CTTGTACCGGT---TTGCATACGCGGTAAAGCGGTAGCGGAAGG 980
Qy 168 GluGluGluGluGluMetGluArgMetLysArgProLysProLysIleLeuGlnThr 187
Db 979 AGAGAACAGAGAGGAGCGCTGCAGAGCTGGAGGTTCAGAGCTCGGCCAGTCCAGGTC 920
Qy 188 ArgArgPro---GluTyrThrProLeuHis 196
```

```

? TOPLOGY: linear
? MOLECULE TYPE: cDNA to mRNA
? HYPOTHETICAL: NO
? ANTI-SENSE: NO
? ORIGINAL SOURCE:
? ORGANISM: Mus musculus
? FEATURE:
? NAME/KEY: repeat_unit
? LOCATION: 1..407
? OTHER INFORMATION: /rpt_type="terminal"
? OTHER INFORMATION: /note="5'UTR"
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: 7471..7476
? OTHER INFORMATION: /function="PolyA_signal"
? FEATURE:
? NAME/KEY: repeat_unit
? LOCATION: 7368..7493
? OTHER INFORMATION: /rpt_type="terminal"
? OTHER INFORMATION: /note="3'UTR"
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: 408..7367
? OTHER INFORMATION: /product="Coagulation Factor VIII"
? PUBLICATION INFORMATION:
? AUTHORS: Eldor, F.
? AUTHORS: Lachin, D.
? TITLE: Sequence of the Murine Factor VIII cDNA.
? JOURNAL: Genomics
? VOLUME: 16
? PAGES: 374-379
? DATE: 1993
? RELEVANT RESIDUES IN SEQ ID NO: 5: FROM 1 TO 7476
PCT-US94-13200-5

Alignment Scores:
Pred. No.: 29.4 Length: 7493
Score: 79.50 Matches: 56
Percent Similarity: 39.19% Conservative: 31
Best Local Similarity: 25.23% Mismatches: 83
Query Match: 7.59% Indels: 52
DB: 5 Gaps: 10

US-09-782-953-3 (1-198) x PCT-US94-13200-5 (1-7493)
QY 15 AlaCySHisLeuAspProArg-----ValPheValAspGlyLeuCysArgAlaIys 31
Db 188 GCCTCCTTCCTCAAGTCCCAAGAAGAACTATCTTCAGATCTGTGGTGAATGCCTAC 247
QY 32 PheGluSer-----LeuPheArgThrTyAsp 40
Db 248 TTTCACCTCACAGTAGATAATAACTCCAGAAAATCTCTGCAGAAAATATTAGACTTTTAC 307
QY 41 LysAspTrpTrpPheGlnTyrPheLys-----SerPheLysArgValArgIleasn 57
Db 308 TAAATCATTAACATTTCTTTTGTTCTTAAAGCTAAAGTTATTTTAAAGAAAGACTTAAAT 367
QY 58 -----PheSerAsnProLeuSerAlaIleAlaAspAlaArgLeuArg 70
Db 368 TTTCATTTCTTACCTTGAACAATTTCTTCTAGTAATAAAGCCATGCAGAAATAGCACTTTCGC 427
QY 71 LeuHisLysThrGlu-----PheLeu-GlyLysGluMetLysLeuTyrPheAlaGI 87
Db 428 TTGCTTCTTCTGAGACCCTTTCAATTTCTGCTCAAGGCCATCAGAAAGATACCTACCTGG 487
QY 87 nThreU-----HisIleGlySerSerHisLeuAlaProAsnProas 102
Db 488 TGCAAGTGAATTCCTCGGAACATATATTCAGAGATCTGCTCAGTGTGCTCATACAGA 547
QY 102 pLyseGlnPheLeuIleSerProProAlaSerProProValGlyTrpLysGlnValGluas 122
Db 548 CTCAGAGATTTCTTCTCAAGATGTCAACATCTTTTCCATTCACCAACCTCATCATGATATAA 607

```

```

Qy      122 palathrProvalIleaeNTYr---AspLeuLeuTYAlAlIseerYs----- 137
      |||
      :|:
      :|:
Db      608 AAGAGCTGCTTTTGAGAGTCAAGAGACCACTTTTCAACATTGGCCAAAGCCAGCCACC 667
      |||
Qy      138 -----LeuGIyProGIyGluYsTYrGIuLeuHIs----- 147
      |||
      |||
      |||
Db      668 CTGATGGGTTGCTAGAGTCTTACCATTTGGACTGAGGTTCAATGACACAGTGGTCAATTAC 727
      |||
Qy      148 -----AlaAlathrAspProthrProserValValValHisValYsGlu---Se 163
      |||
      |||
      |||
Db      728 ACTTAAACAATGCGTCTTCTCATCTTCAGTCTTCATGCTGGTGGTGTGCTTACTGGAA 787
      |||
      |||
      |||
Qy      163 rAspGIuGluBngIuGIuGIuGIuGluMetGluArgMetLeuArgProIyProIy 183
      |||
      |||
      |||
Db      788 ACCTCTCTAGGAGAGATGAATATGAAATCAACAAGCCAAATGAGAGGAAGATGATTA 847
      |||
      |||
      |||
Qy      183 eile 184
      ||:
      ||:
Db      848 AGTT 851

RESULT 11
US-08-261-465-3/C
; Sequence 3, Application US/08261465
; Patent No. 5612191
GENERAL INFORMATION:
APPLICANT: BRIGGS, Steven P.
APPLICANT: BENSEN, Robert J.
TITLE OF INVENTION: PLANT GENES AFFECTING GIBBERELLIC ACID
TITLE OF INVENTION: BIOSYNTHESIS
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/261,465
FILING DATE: 17-JUN-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 33229/213/PIH1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2784 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-261-465-3

Alignment Scores:
Pred. No.: 7.75 Length: 2784
Percent Similarity: 79.00 Matches: 49
Conservative: 31.78% Conservative: 26
Best Local Similarity: 20.76% Mismatches: 79
Query Match: 7.54% Indels: 82
Gaps: 11

US-09-782-953-3 (1-198) x US-08-261-465-3 (1-2784)
Qy      19 AspProArgValPheValaIePglYLeuCYaRGAlaLYePheGluSerLeuPheArgThr 38

```



```

Db 428 TTGCTTCTTCTGAGCTTTTCAATTTCTGCTAGTCCATGACGAAAGATACTACTTGG 487
Qy 87 nhrleu-----HisIleGlySerSerHisLeuAlaProProAnProAs 102
Db 488 TGCAGTGAATTGCTGTGAACATATTCTAGAGTATGATCTGTGCTGTGCTGATACAGA 547
Qy 102 ptysglnpheleuIleSerProProAlaSerProProValGlyTrrplyslnValGluAs 122
Db 548 CTGAAGATTTCTCTTGAAGATGTCATACATCTTTTCCATTGACACCTTCATCATGTATTA 607
Qy 122 pAlaThrProValIleSerTyr---AapLeuLeuTyrAlaIleSerIys----- 137
Db 608 AAGAGCTGTGTGTGTAGAGTACAGAGCCAGCTTTTCAACATGTGCCAAGCCAGCCACC 667
Qy 138 -----LeuGlyProGlyGlyIuTyGlyLeuHis----- 147
Db 668 CTGAGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 727
Qy 148 -----AlaAlaThrAspProThrProSerValValAlaHisValCysGlu---Se 163
Db 728 ACTTAAACATGAGCTTCTATCTCTGAGCTTTCATGCTGTGTGTGTGTGTGTGTGTGT 787
Qy 163 rAapGlnIuAsnGlnIuGlnIuGlnIuMetGluArgMetLysArgProIysProIy 183
Db 788 AGCTTCTGAGGAGATGATATGATGATGATGATGATGATGATGATGATGATGATGATTA 847
Qy 183 gile 184
Db 848 AGTT 851

RESULT 8
US-09-037-601-5
Sequence 5, Application US/09037601
Patent No. 6180371
GENERAL INFORMATION:
APPLICANT: Lollar, John S.
TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: USA
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/037,601
FILING DATE: 26-JUN-1996
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US94/13200
FILING DATE: 15-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/212,133
FILING DATE: 11-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/864,004
FILING DATE: 07-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Feiber, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 75-95F
TELECOMMUNICATION INFORMATION:
TELEPHONE: 303/499-8080
TELEFAX: 303/499-8089
INFORMATION FOR SEQ. ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 7493 base pairs

```

```

TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: not relevant
MOLECULE TYPE: cDNA to mRNA
HYPOTHEICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE: Mus musculus
FEATURE:
NAME/KEY: repeat unit
LOCATION: 1..407
OTHER INFORMATION: /rpt_type="terminal"
OTHER INFORMATION: /note="5' UTR"
FEATURE:
NAME/KEY: misc feature
LOCATION: 7471..7476
OTHER INFORMATION: /function="polya signal"
FEATURE:
NAME/KEY: repeat unit
LOCATION: 7368..7493
OTHER INFORMATION: /rpt_type="terminal"
OTHER INFORMATION: /note="3' UTR"
FEATURE:
NAME/KEY: misc feature
LOCATION: 408..7367
OTHER INFORMATION: /product="coagulation factor VIII"
PUBLICATION INFORMATION:
AUTHORS: Laskich, D.
AUTHORS: Elder, F.
AUTHORS: Gitshier, J.
TITLE: Sequence of the murine Factor VIII cDNA
Patent No. 6180371
JOURNAL: Genomics
VOLUME: 16
PAGES: 374-379
DATE: 1993
US-09-037-601-5

Alignment Scores:
Pred. No.: 29.4 Length: 7493
Score: 79.50 Matches: 56
Percent Similarity: 39.19% Conservative: 31
Best Local Similarity: 25.23% Mismatches: 83
Query Match: 7.59% Indels: 52
DB: Gaps: 10

US-09-782-953-3 (1-198) x US-09-037-601-5 (1-7493)
Qy 15 AlaCysHisLeuAspProArg-----ValPheValAspGlyLeuCyArgAlaLys 31
Db 188 GCCTGCTTCTCTAGTACCCAAAGAAAGTAATCCTTCAGATCTGTTTGTGTAATGCTAC 247
Qy 32 PheGlnSer-----LeuPheArgThrTyrAsp 40
Db 248 TTTCACCTCAGATGATTAACCTCCAGAAATCTCTGCMAAATATTGAGACTTTTAC 307
Qy 41 LysAspThrThrPheGlnTyrPheLys-----SerPheLysArgValArgIleAsn 57
Db 308 TAAATCATTAACATTTCTTTGTTGTTCTTAAGCTTAAGTTATTTTGAAGAAGATTAAAT 367
Qy 58 -----PheSerAsnProLeuSerAlaAlaAspAlaArgLeuArg 70
Db 368 TTTCATTTCTTAGTTGAAACATTTTCTAGTATTAATAAGCCATGCAAAATGCACTTGGC 427
Qy 71 LeuHisLysThrGlu-----PheLeu-GlyLysGlnMetLysLeuTyrPheAlaGln 87
Db 428 TTGCTTCTTCTGAGCTTTTCAATTTCTGCTCTAGTGCATCAGAAAGATACCTTGG 487
Qy 87 nhrleu-----HisIleGlySerSerHisLeuAlaProProAnProAs 102
Db 488 TGCAGTGAATTGCTGTGAACATATTCTAGAGTATGATCTGTGCTGTGCTGATACAGA 547
Qy 102 ptysglnpheleuIleSerProProAlaSerProProValGlyTrrplyslnValGluAs 122

```

```
QY 58 -----PheSerAsnProLeuSerAlaAlaAspAlaArgLeuArg 70
Db 368 TTTCAATTTCTTTAGTTGAACATTTCTAGTAATAAAGCCATGCAATAGCACTCTTCGC 427
QY 71 LeuHisLysThrGlu-----PheLeu-GlyLysGluMetLysLeuTyrPheAlaG1 87
Db 428 TTGCTTCTTTCTGAGCCTTTTCAATTTCTGCTAGTGCATCAGAGAGTACTACTTGG 487
QY 87 nThrLeu-----HisIleGlySerSerHisLeuAlaProProAsnProAs 102
Db 488 TGCAGTGGAAATGTCCTGGAACTATATTCAGAGTGCATCTGCTAGTGTGCTGCATACAGA 547
QY 102 pLysGlnPheLeuIleSerProProAlaSerProProValGlyTyrPlyLysGlnValGluAs 122
Db 548 CTCAGAGATTTCTTCTAGATGTCAGAGTGCATCTTTCCATTCACACACCTCCATCATGTATA 607
QY 122 pAlaThrProValIleAsnTyr---AspLeuLeuTyrAlaIleSerLys----- 137
Db 608 AAGACTGTGTGTTCTAGAGTACAGGACCAAGCTTTTCAACATTCGCAAGCCAGCCAGCCACC 667
QY 138 -----LeuGlyProGlyGluLysTyrGluLeuHis----- 147
Db 668 CTGATGGGTTTGTAGTCTTACCATTTGGACTGGCTTCATGACACAGTGGTTCATTAC 727
QY 148 -----AlaAlaThrAspProThrProSerValValValHisValCysGlu---Se 163
Db 728 ACTTAAAAACATGCTTCTCATCTGTCAGTCTTTCATGCTGTTGGTGTGCTTACTGGAA 787
QY 163 rAspGlnGluAsnGluGluGluGluGluMetGluArgMetLysArgProLysProLy 183
Db 788 AGCTTCTGAGGAGATGAATATAGATCAGACAGCAAGCCAAATGAGGAAGGAAGATGTATA 847
QY 183 sile 184
Db 848 AGTT 851

RESULT 7
US-08-670-707A-5
; Sequence 5, Application US/08670707A
; Patent No. 5859204
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S.
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/670,707A
; FILING DATE: 26-JUN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US94/13200
; FILING DATE: 15-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/212,133
; FILING DATE: 11-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/864,004
; FILING DATE: 07-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Greenlee, Lorraine L.
; REGISTRATION NUMBER: 27,894

; REFERENCE/DOCKET NUMBER: 75-95F
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 303/499-8080
; TELEFAX: 303/499-8089
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7493 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: not relevant
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: repeat unit
; LOCATION: 1..407
; OTHER INFORMATION: /rpt_type= "terminal"
; OTHER INFORMATION: /note= "5' UTR"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 7471..7476
; OTHER INFORMATION: /function= "polyA signal"
; FEATURE:
; NAME/KEY: repeat unit
; LOCATION: 7368..7493
; OTHER INFORMATION: /rpt_type= "terminal"
; OTHER INFORMATION: /note= "3' UTR"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 408..7367
; OTHER INFORMATION: /product= "coagulation factor VIII"
; PUBLICATION INFORMATION:
; AUTHORS: Elder, F.
; AUTHORS: Gitschier, J.
; TITLE: Sequence of the murine Factor VIII cDNA
; Patent No. 5859204
; JOURNAL: Genomics
; VOLUME: 16
; PAGES: 374-379
; DATE: 1993
; US-08-670-707A-5

Alignment Scores:
Pred. No.: 29.4 Length: 7493
Score: 79.50 Matches: 56
Percent Similarity: 39.19% Conservative: 31
Best Local Similarity: 25.23% Mismatches: 83
Query Match: 7.59% Indels: 52
DB: 2 Gaps: 10

US-09-782-953-3 (1-198) x US-08-670-707A-5 (1-7493)
QY 15 AlaCysHisLeuAspProArg-----ValPheValAspGlyLeuCysArgAlaLys 31
Db 188 GCCTGCTTCTTCTACTACTACCACAAAGGAAGTAATCCTTCAGATCTGTTTGTGTAATGCTAC 247
QY 32 PheGluSer-----LeuPheArgThrTyrAsp 40
Db 248 TTTCACTCACAGTAGATAAACTTCCAGAAATCTCTGCAAAATATTTAGGACTTTTAC 307
QY 41 LysAspThrThrPheGlnTyrPheLys-----SerPheLysArgValArgIleAsn 57
Db 308 TAAATCATTAACATTTCTTTTGTAAAGCTAAAGTTATTTTAGAGAGAGAGTTAAAT 367
QY 58 -----PheSerAsnProLeuSerAlaAlaAspAlaArgLeuArg 70
Db 368 TTTCAATTTCTTTAGTTGAACATTTTCTAGTAATAAAGCCATGCAATAGCACTCTTCGC 427
QY 71 LeuHisLysThrGlu-----PheLeu-GlyLysGluMetLysLeuTyrPheAlaG1 87
```


NAME: Forman, David S.
REGISTRATION NUMBER: 33,694
REFERENCE/DOCKET NUMBER: 02481.1321-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4000
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3728 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: Mus musculus
STRAIN: osteoblastic cell line MC3T3E1
FEATURE:
NAME/KEY: CDS
LOCATION: 69..3452
US-08-111-939-1

Alignment Scores:
Pred. No.: 3728
Score: 81.00
Percent Similarity: 38.9%
Best Local Similarity: 20.7%
Query Match: 7.73%
Indels: 60
Gaps: 8
DB:

US-09-782-953-3 (1-198) x US-08-111-939-1 (1-3728)

QY 50 SerPheLysArgValArg-----ile 56
Db 2997 AACTGGAGCGCATTCGGGAGATCTTGCTATGACGGGAACCGTCCCATTCCTCGGAGTT 3056
QY 57 AsnPhSerAsnProLeuSerAlaLaAspAlaArgLeuArgLeuHisLysThrCluPhe 76
Db 3057 GACCCTTCACGACCCATGATGACCCCCCAGCAGCGGCATGACGACGCCGTCTACAGTAC 3116
QY 77 -----LeuGlyLysGluMetLysLeuTyrrPheAlaGlnThrLeuHisIleGlySer 93
Db 3117 CGGCTCGCATGAGGAACAGATGCACCTG-----CGTCGCCTCAATTCTACCGCA 3167
QY 94 SerHisLeuAlaProProAsnProAspLysGlnPheLeuIleSerProProAlaSerPro 113
Db 3168 GGCCCTGCCACAAGCCCACCTCT-----GCCCTTATGCTCCCTCCCTTCCCT 3215
QY 114 -----ProValGlyTrpLys 118
Db 3216 ACACAGCCATTACCTTGGGCGCTGGAGTTCTACCACTACCACTACCACTGGGAG 3275
QY 119 -----GlnValGluAspAlaThrProValIleAsnTyrrAspLeuLeuTyrrAlaIleSer 136
Db 3276 GAGTCAGAGACTGAGACCTATACAGAAGTAGTGACA----- 3311
QY 137 LysLeuGlyProGlyGluLysTyrrGluLeuHisAlaAlaThrAspProThrProSerVal 156
Db 3312 -----GAGTTTGACAGACAGTATGGGACTGAC----- 3338
QY 157 ValValHisValCyGlySerAspGlnGluAsnGluGluGluGluMetGlu 175
Db 3339 ---CTAGAGTGGNAGACATAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3392

RESULT 5
US-08-212-133A-7
; Sequence 7, Application US/08212133A
; Patent No. 5663060
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S.
; APPLICANT: Runge, Marschall S.
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
ADDRESSEE: Kilpatrick & Cody
STREET: 100 Peachtree Street
CITY: Atlanta
STATE: Georgia
COUNTRY: US
ZIP: 30303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/212,133A
FILING DATE: March 11, 1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/864,004
FILING DATE: 07-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: EMU/76677
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-572-6508
TELEFAX: 404-572-6555
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 7493 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: repeat_unit
LOCATION: 1..407
OTHER INFORMATION: /rpt type= "terminal"
OTHER INFORMATION: /note= "5'UTR"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7475..7476
OTHER INFORMATION: /function= "PolyA_signal"
FEATURE:
NAME/KEY: repeat_unit
LOCATION: 7368..7493
OTHER INFORMATION: /rpt type= "terminal"
OTHER INFORMATION: /note= "3'UTR"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 408..7367
OTHER INFORMATION: /product= "Coagulation Factor VIII"
PUBLICATION INFORMATION:
AUTHORS: Elder, F.
AUTHORS: Lekich, D.
AUTHORS: Gitschier, J.
TITLE: Sequence of the Murine Factor VIII cDNA.
Patent No. 5663060
JOURNAL: Genomics
VOLUME: 16
PAGES: 374-379
DATE: 1993
RELEVANT RESIDUES IN SEQ ID NO: 7: FROM 1 TO 7476
US-08-212-133A-7

Alignment Scores:
Pred. No.: 29, 4
Score: 79.50
Percent Similarity: 39.1%
Best Local Similarity: 25.23%
Query Match: 7.59%
Length: 7493
Matches: 56
Conservative: 31
Mismatches: 83
Indels: 52

Percent Similarity: 37.02% Conservative: 28
 Best Local Similarity: 23.56% Mismatches: 74
 Query Match: 8.73% Indels: 57
 DB: 5 Gaps: 9

US-09-782-953-3 (1-198) x PCT-US94-09752-1 (1-1820)

```

QY 2 GUGUVALAARLEUGLHARLEUPROSERALATHRIEALCYHIELEUAP----- 19
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 92 GAAGAGCTTCAAGCTTGATATTTCTTGCTGAACCTTACAMACACCTTGACGAGC 151
QY 20 -----ProxValPheValSer 25
DB 132 AGCAATGAGCGCCAGACATAGCTCCATCCAGACGGATTAAAGATGATCTACAGC 211
QY 26 ---GlyLeuCyArgAlaIylarPheGluSerLeuPheArgThyTyraPlySerThr 44
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 212 ATGACATGTGCTACGGATGATGGAGCCTGTTTCCGAGTGGCTCCCGGAGCCCTC 271
QY 45 PheGlnTyrrPheIylSerPheIylarValAlarGlieAenPheSerAenProleuSerAla 64
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 272 TTCGCC-----AGCCAGATGATCGGTAT 295
QY 65 AAlaerAlaArgLeuArgLeuHisIylSerThyGluPheLeuGlyGluMetIylLeuTy 84
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 296 GCCGAT-----CTTATGACAGCATTTTCATCAT------TTGCTGAT 334
QY 85 PheAlaGlnThrLeuHisIleGlySerSerHisIylarProProAenPro----- 101
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 335 TACCATCATGACCTCTTCAGAGCTGCCACGTTTGATGCTCCATGACATCAACGCTG 394
QY 102 -----AspIylGlnPheLeuIleSerProProAlaSer-----Pro 113
DB 395 GAACACACATGATGATATCAATGAATGAGATCCCTCCGACACGAGAACCGCAGC 454
QY 114 ProValGlyTyrrPylGlnIylarAlaThrProValIleAenTyraPlySerThr 133
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 455 TCCGTGGATTTCAAGAT-----ACCGATCAAGCGGATCACTGACACGG 502
QY 134 AAlaIleSerIylLeuGlyProGlyGlyIylTyrrGluLeuHisIalAlaThrAspProThr 153
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 503 TCGATTACTGAGATTAAACCCCACTCTCCACATG-----GCC 544
QY 154 ProSerValIylAlaHisIylarCygluSerArgGlnGluangIugIugIugIugIu 173
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 545 CCCCAAGAAATTCAGCATTCCTCATGATGAAGATGATGAAGAGAGAGAGAGAGAG 604
QY 174 MetGluArgMetIylarGProIyls 181
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 605 GAAGATGTGCTGGAACCAAAA 628

```

RESULT 3

US-09-234-332-5/c
 ; Sequence 5, Application US/09234332A
 ; Patent No. 6087168
 ; GENERAL INFORMATION:
 ; APPLICANT: Cedars-Sinai Medical Center
 ; APPLICANT: Michel F. Levesque, M.D.
 ; APPLICANT: Thomas Neuman, Ph.D.
 ; TITLE OF INVENTION: CONVERSION OF NON-NEURONAL CELLS INTO
 ; FILE REFERENCE: P07 41494
 ; CURRENT APPLICATION NUMBER: US/09/234,332A
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: Paeseq for Windows Version 3.0
 ; SEQ ID NO 5
 ; LENGTH: 3138
 ; TYPE: DNA
 ; ORGANISM: Human
 ; FEATURE:
 ; NAME/KEY: gene
 ; LOCATION: (0)...(0)

OTHER INFORMATION: ZIC 1 Protein gene; Genbank Accession D76435
 US-09-234-332-5

Alignment Scores:

Pred. No.:	Score:	Length:	Matches:	Mismatches:	Indels:	Gaps:
5.2	81.00	3138	37	12	50	22
Percent Similarity:	40.50%					
Best Local Similarity:	30.58%					
Query Match:	7.73%					
DB:	3					5

US-09-782-953-3 (1-198) x US-09-234-332-5 (1-3138)

```

QY 86 AAlaGlnThrLeuHisIleGlySerSerHisIylarProProAenProArlyGlnPhe 105
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 950 GCCAGCTGCTGCAAGACTGGGGTTGAGC---TTGAAGCGCCCATCGCTCGCGGAACGG 894
QY 106 LeuIleSerProPro-----AlaSerProProValGlyTyrrGlnIyl 120
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 893 TTGATGCCACAGCCCAAGCTCGTCCGCCAGCTCGCCCGGAGTGTGCGGAGCGG 834
QY 121 ---GluSerAlaThrProValIleAenTyraPlySerThrAlaIleSerIylLeuGly 139
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 833 CCAAGGTGCTACGCGCGATCGCTGGTAC-----TGGGGG 798
QY 140 ProGlyGlyTyrrGluLeuHisIalAlaThrAspProThrProSerVal-ValIylHis 159
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 797 CCGGCGTCCAGAGCATCGTGTGCTCGGGGCGAGACCCCGGCTCCCCCCCCCA 738
QY 159 ValCygluSerArgGlnGluangIugIugIugIugIugIugIugIugIugIugI 179
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 737 CCTTCGACACCGAGGAGGAAATCGGAGGAGAGAGAGAGAGAGAGAGAGAGAG 678
QY 179 rGProIylsPylIleIleGlnThrArgArProGluTyrrThrProIleHisLeuSer 198
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 677 GAGAGAG-----AAGAGAGAGAGAGGTGATGTCACCACTCGCG 634

```

RESULT 4

US-08-111-939-1
 ; Sequence 1, Application US/08111939
 ; Patent No. 5460951
 ; GENERAL INFORMATION:
 ; APPLICANT: Kawai, Shinji
 ; APPLICANT: Takeshita, Sunao
 ; APPLICANT: Okazaki, Makoto
 ; APPLICANT: Amano, Egon
 ; TITLE OF INVENTION: Bone-Related Carboxypeptidase-Like
 ; NUMBER OF SEQUENCES: 27
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
 ; STREET: 1300 I Street, N.W.
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20005-3315
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/111,939
 ; FILING DATE: 26-AUG-1993
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 324033/92
 ; FILING DATE: 03-DEC-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 230029/92
 ; FILING DATE: 28-AUG-1992
 ; ATTORNEY/AGENT INFORMATION:

Sun Dec 15 08:38:14 2002

REGISTRATION NUMBER: 33,778
REFERENCE/DOCKET NUMBER: U010815-9
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 708-1800
TELEFAX: (212) 246-8959
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2174 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: cDNA for mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE: human
ORGANISM: Homo sapiens
DEVELOPMENTAL STAGE: foetal
TISSUE TYPE: Brain
IMMEDIATE SOURCE: gene library of cDNA
LIBRARY: gene library of cDNA from foetal
CLONE: BC-17.8-1 and BC-17.8-2
POSITION IN GENOME: chromosome 21/YAC 72H9
CHROMOSOME/SEGMENT: chromosome 21/q22.1-q22.2
MAP POSITION: 21q22.1-q22.2
FEATURE:
NAME/KEY: cDNA for mRNA, BC-17.8
LOCATION: 1..2174
FEATURE:
NAME/KEY: untranslated 5'
LOCATION: 1..48
FEATURE:
NAME/KEY: coding sequence
LOCATION: 49..560
OTHER INFORMATION: Down Syndrome critical
FEATURE:
NAME/KEY: DSCR1
LOCATION: 1..171
IDENTIFICATION METHOD: translation of the
OTHER INFORMATION: - deduced protein
OTHER INFORMATION: - proline-rich protein domains
OTHER INFORMATION: - glutamic acid-rich protein domains
OTHER INFORMATION: - leucine/phenylalanine-rich protein domains
FEATURE:
NAME/KEY: untranslated 3'
LOCATION: 564..2174
FEATURE:
NAME/KEY: 2 poly (A)
LOCATION: 1541..1546 AND 2132..2137
US-08-665-040-1

Alignment Scores:
Pred. No.: 6,31e-96 Length: 2174
Score: 844.50 Matches: 161
Percent Similarity: 97.63% Conservative: 4
Best Local Similarity: 95.27% Mismatches: 3
Query Match: 80.58% Indels: 1
DB: 2 Gaps: 1

US-09-782-953-3 (1-198) x US-08-665-040-1 (1-2174)

QY 30 AlaLysPheGluSerLeuPheArgThrTyrAspLysAspThrPheGlnTyrPheLys 49
Db 58 GCCAAATTTGAGTCCCTCTTTAGGACGTATGACAGGACATCACCCTTCAGTATTTTAA 117
QY 50 SerPheLysArgValArgLeuPheSerAsnProLeuSerAlaAlaAspAlaArgLeu 69
Db 118 AGCTTCAAAACAGTCAAGATTAATCTTCAGCAACCCCTCTCCGACGACATGCCAGGCTC 177
QY 70 ArgLeuHisLysThrGluPheLeuGlyLysGluMetLysLeuTyrPheAlaGlnThrLeu 89
Db 178 CAGCTGCAATGAGTCTGAGTTTCTGGGAAAGAAATGAAGTTATATTTGCTCAGACCTTA 237

QY 90 HisLleGlySerSerHisLeuAlaProProAsnProAspLysGlnPheLeuIleSerPro 105
Db 238 CACATAGGAAGCTCACACCTGGCTCCGCCAAATCCAGACAGCAGTTTCTGATCTCCCT 297
QY 110 ProAlaSerProProValGlyTyrPheGlnValGluAspAlaThrProValIleAsnTyr 129
Db 298 CCGGCTCTCCGCCAGTGGGATGGAAACAACACTGGAGATGGACCCAGTCTATAACTAT 357
QY 130 AspLeuLeuTyrAlaIleSerLysLeuGlyProGlyGluLysTyrGluLeuHisAlaAla 149
Db 358 GATCTCTTATATGCCATCTCCAAAGCTGGGCCAGGGGAAAGATATGAATTCACGCAGCG 417
QY 150 ThrAspProThrProSerValValHisValCysGluSerAspGlnGluAsnGluGlu 169
Db 418 ACTGACACCACTCCACGCGTGGTGTCCATGTATGTGAGAGTGAAGAG---AAGGAG 474
QY 170 GluGluGluGluMetGluArgMetLysArgProLysPheLysIleIleGlnThrArgArg 189
Db 475 GAAGAGAGAGAAATGGAAAGAAATGAGGACACTTACGCAAAAATATCCAGACCAGGAGG 534
QY 190 ProGluTyrThrProIleHisLeuSer 198
Db 535 CCGGAGTACACGCGGATCCACCTCAGC 561

RESULT 2

PCT-US94-09752-1
Sequence 1, Application PC/TUS9409752
GENERAL INFORMATION:
APPLICANT: David S. Strayer and Avinash Chander
TITLE OF INVENTION: Compositions and Methods for
Targeting Cells and Modulating Pulmonary Surfactant Secretion
TITLE OF INVENTION: Targeting Cells and Modulating Pulmonary Surfactant Secretion
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESS: Jane Massey Licata, Esq.
STREET: 210 Lake Drive East, Suite 201
CITY: Cherry Hill
STATE: NJ
COUNTRY: USA
ZIP: 08002
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb
MEDIUM TYPE: STORAGE
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/09752
FILING DATE: Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/176,218
FILING DATE: December 30, 1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/114,951
FILING DATE: August 31, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Jane Massey Licata
REGISTRATION NUMBER: 32,257
REFERENCE/DOCKET NUMBER: JEFF-0042
TELECOMMUNICATION INFORMATION:
TELEPHONE: (609) 779-2400
TELEFAX: (609) 779-8488
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1820
TYPE: nucleic acid
STRANDEDNESS: single stranded
TOPOLOGY: linear
PCT-US94-09752-1

Alignment Scores:
Pred. No.: 0.112 Length: 1820
Score: 91.50 Matches: 49

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 14, 2002, 20:00:06 ; Search time 30.0506 Seconds

(without alignments)
2020.659 Million cell updates/sec

Title: US-09-782-953-3

Perfect score: 1048

Sequence: 1 MEEVDLQDLPSATTACHLDP.....RPRKXIQTRRPEYTHLS 198

Scoring table:

BLOSUM62	Xgapex 10.0	Xgapex 0.5
Ygapex 10.0	Ygapex 0.5	
Fgapex 6.0	Fgapex 7.0	
Delop 6.0	Delop 7.0	

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing filter 45 summaries

Command line parameters:

-MODEL=frame_plus_p2n.model -DEV=x1h
-Q/cg2_1/USPTO.epool/US09782953/runat.11122002.114430.17377/app.query.fasta.1.1173
-DB=Issued_Patents_NA -OPMT=fastcap -SUFFIX=trn1 -MINMATCH=0.1 -DOORL=0
-LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=50 -ALIGN=15
-MODE=LOCAL -OUTPMT=pc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-NOR=US09782953 @CGN_1.1.29 @runat.11122002.114430.17377 -NCPU=6 -ICPU=3
-NO_XLXRY -NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMOUT=120
-MAEN_TIMOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Issued Patents NA:*

- 1: /cg2_6/prodata/1/ina/5A.COMB.seq:*
- 2: /cg2_6/prodata/1/ina/5B.COMB.seq:*
- 3: /cg2_6/prodata/1/ina/6A.COMB.seq:*
- 4: /cg2_6/prodata/1/ina/6B.COMB.seq:*
- 5: /cg2_6/prodata/1/ina/PCUS.COMB.seq:*
- 6: /cg2_6/prodata/1/ina/backfillseq1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	844.5	80.6	2174	2	US-08-665-040-1
2	91.5	8.7	1820	5	PCT-US94-09752-1
3	81	7.7	3138	3	US-09-234-332-5
4	81	7.7	3728	1	US-08-111-939-1
5	79.5	7.6	7493	1	US-08-212-133A-7
6	79.5	7.6	7493	1	US-08-474-503-5
7	79.5	7.6	7493	2	US-08-670-707A-5
8	79.5	7.6	7493	4	US-09-037-601-5
9	79.5	7.6	7493	4	US-09-315-179-5
10	79.5	7.6	7493	5	PCT-US94-13200-5
11	79	7.5	2784	1	US-08-261-465-3
12	79	7.5	2784	1	US-08-405-254-7

13	79	7.5	3429	1	US-08-097-997A-10	Sequence 10, Appl
14	79	7.5	3429	3	US-08-665-574C-10	Sequence 10, Appl
15	79	7.5	3429	4	US-08-946-994-10	Sequence 10, Appl
16	79	7.5	4234	1	US-08-446-038B-1	Sequence 1, Appl
17	79	7.5	4234	1	US-08-446-010B-1	Sequence 1, Appl
18	79	7.5	4234	1	US-08-805-445-1	Sequence 1, Appl
19	79	7.5	4234	2	US-08-064-067D-1	Sequence 1, Appl
20	79	7.5	4234	2	US-09-066-208-1	Sequence 1, Appl
21	79	7.5	36651	4	US-09-738-894A-3	Sequence 3, Appl
22	78.5	7.5	1875	4	US-09-877-730-25	Sequence 25, Appl
23	78.5	7.5	1875	4	US-09-877-730-23	Sequence 23, Appl
24	78.5	7.5	1877	4	US-09-877-730-22	Sequence 22, Appl
25	78.5	7.5	2139	4	US-09-877-730-21	Sequence 21, Appl
26	78.5	7.5	2382	4	US-09-877-730-20	Sequence 20, Appl
27	78.5	7.5	2481	4	US-09-877-730-19	Sequence 19, Appl
28	78.5	7.5	2715	4	US-09-877-730-18	Sequence 18, Appl
29	78.5	7.5	2724	4	US-09-877-730-17	Sequence 17, Appl
30	78.5	7.5	2958	4	US-09-877-730-16	Sequence 16, Appl
31	78.5	7.5	2976	4	US-09-877-730-15	Sequence 15, Appl
32	78.5	7.5	3210	4	US-09-877-730-14	Sequence 14, Appl
33	78.5	7.5	3219	4	US-09-877-730-13	Sequence 13, Appl
34	78.5	7.5	3453	4	US-09-877-730-12	Sequence 12, Appl
35	78.5	7.5	3874	4	US-09-877-730-11	Sequence 11, Appl
36	78.5	7.5	6171	1	US-08-459-568-1	Sequence 1, Appl
37	78.5	7.5	6171	2	US-08-399-411-1	Sequence 2, Appl
38	78.5	7.5	6171	3	US-08-516-859A-1	Sequence 3, Appl
39	78.5	7.5	6171	4	US-09-586-472-1	Sequence 4, Appl
40	78.5	7.5	6171	4	US-09-528-706-1	Sequence 4, Appl
41	78	7.4	1913	2	US-08-588-258B-41	Sequence 41, Appl
42	78	7.4	1913	5	PCT-US96-08295-41	Sequence 6, Appl
43	77	7.3	1086	2	US-08-985-090-6	Sequence 3, Appl
44	77	7.3	1086	3	US-09-165-543-33	Sequence 50, Appl
45	77	7.3	1740	4	US-08-796-101-50	

ALIGNMENTS

RESULT 1
US-08-665-040-1
Sequence 1, Application US/0865040
Patent No. 5865318
GENERAL INFORMATION:
APPLICANT: ESTIVILL PALLEJA, XAVIER
APPLICANT: FUENTES, JUAN JOSE
APPLICANT: PRITCHARD, MELANIE
TITLE OF INVENTION: A NEW GENE SEQUENCE OF THE
TITLE OF INVENTION: DOWN SYNDROME CRITICAL REGION OF HUMAN CHROMOSOME 21,
TITLE OF INVENTION: IDENTIFIED BY A NEW "ALU-SPICING-PCR" TECHNIQUE,
TITLE OF INVENTION: CODING FOR A PROLINE-RICH PROTEIN (DSCT1) HIGHLY
TITLE OF INVENTION: EXPRESSED IN FOETAL BRAIN AND IN HEART AND METHOD
NUMBER OF SEQUENCES: 12
FOR CHARACTERIZING IT.
CORRESPONDENCE ADDRESS:
ADDRESSEE: LADAS & PARRY
STREET: 26 WEST 61ST STREET
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10023
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORD PERFECT 5.1 FOR DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/665,040
CLASSIFICATION: 435
FILING DATE: JUNE 7, 1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: ES P9501140
FILING DATE: JUNE 7, 1995
ATTORNEY/AGENT INFORMATION:
NAME: JANET I. CORD

GenCore version 5.1.3
Copyright (C) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 14, 2002, 20:02:01 ; Search time 1822.74 Seconds
(without alignments)
2731.177 Million cell updates/sec

Title: US-09-782-953-3
Perfect score: 1048
Sequence: 1 MSEVDLQDLPSTAIACHLDP.....RPKXIIQTRRPEYTPIHLS 198

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delgap 6.0 , Delext 7.0

Searched: 24791104 seqs, 12571243825 residues
Total number of hits satisfying chosen parameters: 49582208

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.epool/US09782953/runat_11122002_114430_17408/app_query.fasta.1.1173
-DB=Pending Patents NA Main -QMT-fastap -SUFFIX=nrnp -MINMATCH=0.1 -LOOPCL=0
-LOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS-human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09782953 -CGN 1 1 2457 -runat_11122002_114430_17408 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAB -LARGQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Pending Patents NA Main:
1: /cgn2_6/ptodata/1/pna/PCTUS_COMB.seq.*
2: /cgn2_6/ptodata/1/pna/US06_COMB.seq.*
3: /cgn2_6/ptodata/1/pna/US07_COMB.seq.*
4: /cgn2_6/ptodata/1/pna/US080_COMB.seq.*
5: /cgn2_6/ptodata/1/pna/US081_COMB.seq.*
6: /cgn2_6/ptodata/1/pna/US082_COMB.seq.*
7: /cgn2_6/ptodata/1/pna/US083_COMB.seq.*
8: /cgn2_6/ptodata/1/pna/US084_COMB.seq.*
9: /cgn2_6/ptodata/1/pna/US085_COMB.seq.*
10: /cgn2_6/ptodata/1/pna/US086_COMB.seq.*
11: /cgn2_6/ptodata/1/pna/US087_COMB.seq.*
12: /cgn2_6/ptodata/1/pna/US088_COMB.seq.*
13: /cgn2_6/ptodata/1/pna/US089_COMB.seq.*
14: /cgn2_6/ptodata/1/pna/US090_COMB.seq.*
15: /cgn2_6/ptodata/1/pna/US091_COMB.seq.*
16: /cgn2_6/ptodata/1/pna/US092_COMB.seq.*
17: /cgn2_6/ptodata/1/pna/US093_COMB.seq.*
18: /cgn2_6/ptodata/1/pna/US094_COMB.seq.*
19: /cgn2_6/ptodata/1/pna/US095A_COMB.seq.*
20: /cgn2_6/ptodata/1/pna/US095B_COMB.seq.*
21: /cgn2_6/ptodata/1/pna/US095C_COMB.seq.*
22: /cgn2_6/ptodata/1/pna/US095D_COMB.seq.*
23: /cgn2_6/ptodata/1/pna/US096A_COMB.seq.*
24: /cgn2_6/ptodata/1/pna/US096B_COMB.seq.*
25: /cgn2_6/ptodata/1/pna/US096C_COMB.seq.*
26: /cgn2_6/ptodata/1/pna/US096D_COMB.seq.*
27: /cgn2_6/ptodata/1/pna/US096E_COMB.seq.*
28: /cgn2_6/ptodata/1/pna/US097A_COMB.seq.*

29:	/cgn2_6/ptodata/1/pna/US097B_COMB.seq.*
30:	/cgn2_6/ptodata/1/pna/US097C_COMB.seq.*
31:	/cgn2_6/ptodata/1/pna/US098A_COMB.seq.*
32:	/cgn2_6/ptodata/1/pna/US098B_COMB.seq.*
33:	/cgn2_6/ptodata/1/pna/US098C_COMB.seq.*
34:	/cgn2_6/ptodata/1/pna/US099A_COMB.seq.*
35:	/cgn2_6/ptodata/1/pna/US099B_COMB.seq.*
36:	/cgn2_6/ptodata/1/pna/US099C_COMB.seq.*
37:	/cgn2_6/ptodata/1/pna/US099D_COMB.seq.*
38:	/cgn2_6/ptodata/1/pna/US100A_COMB.seq.*
39:	/cgn2_6/ptodata/1/pna/US100B_COMB.seq.*
40:	/cgn2_6/ptodata/1/pna/US101A_COMB.seq.*
41:	/cgn2_6/ptodata/1/pna/US101B_COMB.seq.*
42:	/cgn2_6/ptodata/1/pna/US102A_COMB.seq.*
43:	/cgn2_6/ptodata/1/pna/US102B_COMB.seq.*
44:	/cgn2_6/ptodata/1/pna/US6000_COMB.seq.*
45:	/cgn2_6/ptodata/1/pna/US6001_COMB.seq.*
46:	/cgn2_6/ptodata/1/pna/US6002_COMB.seq.*
47:	/cgn2_6/ptodata/1/pna/US6003_COMB.seq.*
48:	/cgn2_6/ptodata/1/pna/US6004_COMB.seq.*
49:	/cgn2_6/ptodata/1/pna/US6005_COMB.seq.*
50:	/cgn2_6/ptodata/1/pna/US6006_COMB.seq.*
51:	/cgn2_6/ptodata/1/pna/US6007_COMB.seq.*
52:	/cgn2_6/ptodata/1/pna/US6008_COMB.seq.*
53:	/cgn2_6/ptodata/1/pna/US6009_COMB.seq.*
54:	/cgn2_6/ptodata/1/pna/US6010_COMB.seq.*
55:	/cgn2_6/ptodata/1/pna/US6011_COMB.seq.*
56:	/cgn2_6/ptodata/1/pna/US6012_COMB.seq.*
57:	/cgn2_6/ptodata/1/pna/US6013_COMB.seq.*
58:	/cgn2_6/ptodata/1/pna/US6014_COMB.seq.*
59:	/cgn2_6/ptodata/1/pna/US6015_COMB.seq.*
60:	/cgn2_6/ptodata/1/pna/US6016_COMB.seq.*
61:	/cgn2_6/ptodata/1/pna/US6017_COMB.seq.*
62:	/cgn2_6/ptodata/1/pna/US6018_COMB.seq.*
63:	/cgn2_6/ptodata/1/pna/US6019_COMB.seq.*
64:	/cgn2_6/ptodata/1/pna/US6020_COMB.seq.*
65:	/cgn2_6/ptodata/1/pna/US6021_COMB.seq.*
66:	/cgn2_6/ptodata/1/pna/US6022_COMB.seq.*
67:	/cgn2_6/ptodata/1/pna/US6023_COMB.seq.*
68:	/cgn2_6/ptodata/1/pna/US6024_COMB.seq.*
69:	/cgn2_6/ptodata/1/pna/US6025_COMB.seq.*
70:	/cgn2_6/ptodata/1/pna/US6026_COMB.seq.*
71:	/cgn2_6/ptodata/1/pna/US6027_COMB.seq.*
72:	/cgn2_6/ptodata/1/pna/US6028_COMB.seq.*
73:	/cgn2_6/ptodata/1/pna/US6029_COMB.seq.*
74:	/cgn2_6/ptodata/1/pna/US6030_COMB.seq.*
75:	/cgn2_6/ptodata/1/pna/US6031_COMB.seq.*
76:	/cgn2_6/ptodata/1/pna/US6032_COMB.seq.*
77:	/cgn2_6/ptodata/1/pna/US6033_COMB.seq.*
78:	/cgn2_6/ptodata/1/pna/US6034_COMB.seq.*
79:	/cgn2_6/ptodata/1/pna/US6035_COMB.seq.*
80:	/cgn2_6/ptodata/1/pna/US6036_COMB.seq.*
81:	/cgn2_6/ptodata/1/pna/US6037_COMB.seq.*
82:	/cgn2_6/ptodata/1/pna/US6038_COMB.seq.*
83:	/cgn2_6/ptodata/1/pna/US6039_COMB.seq.*
84:	/cgn2_6/ptodata/1/pna/US6040_COMB.seq.*
85:	/cgn2_6/ptodata/1/pna/US6041_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1048	100.0	597	22	US-09-575-580B-2
2	1048	100.0	597	30	US-09-782-953-2
3	1040	99.2	597	80	US-60-360-207-8520
4	1002.5	95.7	2212	30	US-09-782-953-14
5	1002.5	95.7	2407	40	US-10-104-047-922
6	957.5	91.4	2358	1	PCT-US00-05918-347


```
Db 85 ||||| CCGGCTGTGTGGACCGGCTGTGCGCGCCAAATTGAGTCCCTCTTGGACGATGAC 144
Qy 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60
Db 145 AAGGACATCACTTTCAGATTTTAAAGCTTCAACGAGCTCAAGATTAACCTTCAGCAAC 204
Qy 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
Db 205 CCGTCTCCGCGACGACATGCCAGGCTCCAGCTGCATTAAGACTGATTTCTGGGAAAGGAA 264
Qy 81 MetLysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProProAsn 100
Db 265 ATGAAGTTATATTGTTGCTCAGACCTTACATAGGAAGCTCAACACTGGCTCCGCCAAAT 324
Qy 101 ProAspLysGlnPheLeuIleSerProProAlaSerProValGlyTyrPlyGlnVal 120
Db 325 CCAGACAAACAGTTTCTGATCTCCCTCCGCTCTCCGCCAGTGGAGTGGAAACAGTG 384
Qy 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro 140
Db 385 GAAGATGCCAGCCCGACTCATAACTATGATCTCTTATATGCGCATCTCCAAGCTGGGGCCA 444
Qy 141 GlyLysTyrTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160
Db 445 GGGGAAAAGTATGAATTGACGCGAGCTGACACCACTCCAGCGCTGGTGGTCCATGTA 504
Qy 161 CysGluSerAspGlnGluAsnGluGluGluGluGluMetGluArgMetLysArgPro 180
Db 505 TGTGAGATGATCAAGAG---AAGAGGAAAGAAAGGAAATGGAAAGATGAGAGACT 561
Qy 181 LysProLysIleIleGlnThrArgArgProGluTyrThrProIleHisLeuSer 198
Db 562 AAGCCAAAATTATTCAGACGACGAGGCGCGAGTACACGCCATCACCCTCAGC 615

RESULT 5
US-10-104-047-922
: Sequence 922, Application US/10104047
: GENERAL INFORMATION:
: APPLICANT: HELIX RESEARCH INSTITUTE
: TITLE OF INVENTION: Novel full length cDNA
: FILE REFERENCE: H1-A0105
: CURRENT APPLICATION NUMBER: US/10/104,047
: PRIOR FILING DATE: 2002-03-25
: PRIOR APPLICATION NUMBER:
: NUMBER OF SEQ ID NOS: 4096
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 922
: LENGTH: 2407
: TYPE: DNA
: ORGANISM: Homo sapiens
US-10-104-047-922

Alignment Scores:
Pred. No.: 5.25e-91 Length: 2407
Score: 1002.50 Matches: 191
Percent Similarity: 97.98% Conservative: 3
Best Local Similarity: 96.46% Mismatches: 3
Query Match: 95.66% Indels: 1
DB: 40 Gaps: 1

US-09-782-953-3 (1-198) x US-10-104-047-922 (1-2407)
Qy 1 MetGluGluValAspLeuGlnAspLeuProSerAlaThrIleAlaCysHisLeuAspPro 20
Db 218 ATGGAGGAGGTGGACCTGCGAGACCTGCCAGCGCCACCATGCTGCTTCACCTGAGACCG 277
Qy 21 ArgValPheValAspGlyLeuCysArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40
Db 278 CGCGTGTTCGTGGAGCGGCTGTGCGGGCCAAATTGAGTCCCTCTTAGGACGTATGAC 337
Qy 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60
```

```
Db 338 AAGGACATCACTTTCAGATTTTAAAGCTTCAACGAGCTCAAGATTAACCTTCAGCAAC 397
Qy 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
Db 388 CCGTCTCCGCGACGACGATGCCAGGCTCCAGCTCAATTAAGACTGATTTCTGGGAAAGGAA 457
Qy 81 MetLysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProProAsn 100
Db 458 ATGAAGTTATATTGTTGCTCAGACCTTACATAGGAAGCTCAACACTGGCTCCGCCAAAT 517
Qy 101 ProAspLysGlnPheLeuIleSerProProAlaSerProProValGlyTyrPlyGlnVal 120
Db 518 CCAGACAAACAGTTTCTGATCTCCCTCCGCTCTCCGCCAGTGGAGTGGAAACAACTG 577
Qy 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro 140
Db 578 GAAGATGCCAGCCCGACTCATAACTATGATCTCTTATATGCGCATCTCCAAGCTGGGGCCA 637
Qy 141 GlyLysTyrTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160
Db 638 GGGGAAAAGTATGAATTGACGCGAGCTGACACCACTCCAGCGCTGGTGGTCCATGTA 697
Qy 161 CysGluSerAspGlnGluAsnGluGluGluGluGluMetGluArgMetLysArgPro 180
Db 698 TGTGAGATGATCAAGAG---AAGAGGAAAGAAAGGAAATGGAAAGATGAGAGACT 754
Qy 181 LysProLysIleIleGlnThrArgArgProGluTyrThrProIleHisLeuSer 198
Db 755 AAGCCAAAATTATTCAGACGACGAGGCGCGAGTACACGCCATCACCCTCAGC 808

RESULT 6
PCT-US00-05918-347
: Sequence 347, Application PC/TUS0005918
: GENERAL INFORMATION:
: APPLICANT: Craig Rosen,
: APPLICANT: Steve Ruben
: TITLE OF INVENTION: Human Lung Cancer Associated Gene Sequences and Polypeptides
: FILE REFERENCE: PA104PCT
: CURRENT APPLICATION NUMBER: PCT/US00/05918
: PRIOR FILING DATE: 2000-03-08
: EARLIER APPLICATION NUMBER: 60/124,270
: NUMBER OF SEQ ID NOS: 896
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 347
: LENGTH: 2358
: TYPE: DNA
: ORGANISM: Homo sapiens
PCT-US00-05918-347

Alignment Scores:
Pred. No.: 1.91e-86 Length: 2358
Score: 957.50 Matches: 188
Percent Similarity: 96.46% Conservative: 3
Best Local Similarity: 94.95% Mismatches: 6
Query Match: 91.36% Indels: 2
DB: 1 Gaps: 1

US-09-782-953-3 (1-198) x PCT-US00-05918-347 (1-2358)
Qy 1 MetGluGluValAspLeuGlnAspLeuProSerAlaThrIleAlaCysHisLeuAspPro 20
Db 114 ATGGAGGAGGTGGACCTGCGAGACCTGCCAGCGCCACCATGCTGCTTCACCTGAGACCG 173
Qy 21 ArgValPheValAspGlyLeuCysArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40
Db 174 CGCGTGTTCGTGGAGCGGCTGTGCGGGCCAAATTGAGTCCCTCTTAGGACGTATGAC 233
Qy 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60
Db 234 AAGGACATCACTTTCAGATTTTAAAGCTTCAACGAGCTCAAGATTAACCTTCAGCAAC 293
```

```
Qy 61: ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLeuThrGluPheLeuGlyLysGlu 80
|||
Db 294 CCCTTCCGCGCAGATGCCAGCTCCAGCTGCATAGACTGAGTTTCTGGGAAGGAA 353
|||
Qy 81 MetLysLeuTyrPheAlaGlnThrLeuHisLeuGlySerHisLeuAlaProProAsn 100
|||
Db 354 ATGAAGTTATATTTTGTCTGACAGCTTACACATAGGAAGCTCACACCTGGCTCCGCA-AA 412
|||
Qy 101 ProAspLysGlnPheLeuLeuSerProProAlaSerProProValGlyTyrLysGlnVal 120
|||
Db 413 CCAGAACAGCAGTTTCTGATCTCCCTCCCGCTCTCCGSCAGTGGATGGAACAAGTG 472
|||
Qy 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro 140
|||
Db 473 GAAGATGCGACCCCGATCAATAAATGATCTCTTATATGCCATCTCCAAAGCTGGGGCA 532
|||
Qy 141 GlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160
|||
Db 533 GGGGAAAAGTATGAATTTGCACGCGAGCTGACACCACTCCCGAGCTGGTGGTCCATGTA 592
|||
Qy 161 CysGluSerAspGlnGluAsnGluGluGluMetGluArgMetLysArgPro 180
|||
Db 593 TGTRAGAGTATCAAGAG---AAGGAGGAAGAGGAAATGGAAAGAAATGAGGAGACCT 649
|||
Qy 181 LysProLysIleLeuGlnThrArgArgProGluTyrThrProIleHisLeuSer 198
|||
Db 650 AAGCCAAAATTTATCCAGACGAGGCGGAGTACACGCGGATCCACCTCAGC 703
|||
RESULT 7
US-09-925-302-347
; Sequence 347, Application US/09925302
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: P4104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 347
; LENGTH: 2358
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-302-347
Alignment Scores:
Pred. No.: 1,91e-86 Length: 2358
Score: 957.50 Matches: 188
Percent Similarity: 96.46% Conservative: 3
Best Local Similarity: 94.95% Mismatches: 6
Query Match: 91.36% Indels: 2
DB: 34 Gaps: 1
US-09-782-953-3 (1-198) x US-09-925-302-347 (1-2358)
Qy 1 MetGluGluValAspLeuGlnAspLeuProSerAlaThrIleAlaCysHisLeuAspPro 20
|||
Db 114 ATGAGGAGGTGGACCTGACGACCTGCCAGCGCCACCATCGCTGTCACTGGACCCG 173
|||
Qy 21 ArgValPheValAspGlyLeuCysArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40
|||
Db 174 CCGCGTGTTCGTGGAGCGGCTGTGCGGGGCAAAATTTGAGTCCCTCTTTAGACGTATGAC 233
|||
Qy 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgValArgIleAsnPheSerAsn 60
|||
Db 234 AAGGACATCACCTTTTCAGTATTTTAAGAGCTTCAAAACGAGTCAGATAAATCTTCAGCAAC 293
|||
Qy 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysGlu 80
|||
```

```
Db 294 CCCTTCCGCGCAGATGCCAGCTCCAGCTGCATAGACTGAGTTTCTGGGAAGGAA 353
|||
Qy 81 MetLysLeuTyrPheAlaGlnThrLeuHisLeuGlySerHisLeuAlaProProAsn 100
|||
Db 354 ATGAAGTTATATTTTGTCTGACAGCTTACACATAGGAAGCTCACACCTGGCTCCGCA-AA 412
|||
Qy 101 ProAspLysGlnPheLeuLeuSerProProAlaSerProProValGlyTyrLysGlnVal 120
|||
Db 413 CCAGAACAGCAGTTTCTGATCTCCCTCCCGCTCTCCGSCAGTGGATGGAACAAGTG 472
|||
Qy 121 GluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyPro 140
|||
Db 473 GAAGATGCGACCCCGATCAATAAATGATCTCTTATATGCCATCTCCAAAGCTGGGGCA 532
|||
Qy 141 GlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisVal 160
|||
Db 533 GGGGAAAAGTATGAATTTGCACGCGAGCTGACACCACTCCCGAGCTGGTGGTCCATGTA 592
|||
Qy 161 CysGluSerAspGlnGluAsnGluGluGluMetGluArgMetLysArgPro 180
|||
Db 593 TGTRAGAGTATCAAGAG---AAGGAGGAAGAGGAAATGGAAAGAAATGAGGAGACCT 649
|||
Qy 181 LysProLysIleLeuGlnThrArgArgProGluTyrThrProIleHisLeuSer 198
|||
Db 650 AAGCCAAAATTTATCCAGACGAGGCGGAGTACACGCGGATCCACCTCAGC 703
|||
RESULT 8
US-09-644-868-7429
; Sequence 7429, Application US/09644868
; GENERAL INFORMATION:
; APPLICANT: White, David
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Pan, Yang
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.1168-001
; CURRENT APPLICATION NUMBER: US/09/644,868
; CURRENT FILING DATE: 2000-08-28
; PRIOR APPLICATION NUMBER: 60/151,063
; PRIOR FILING DATE: 1999-08-27
; NUMBER OF SEQ ID NOS: 10075
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7429
; LENGTH: 1615
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1615)
; OTHER INFORMATION: n = A,T,C or G
US-09-644-868-7429
Alignment Scores:
Pred. No.: 3,04e-86 Length: 1615
Score: 953.50 Matches: 188
Percent Similarity: 95.50% Conservative: 3
Best Local Similarity: 94.00% Mismatches: 6
Query Match: 90.98% Indels: 3
DB: 25 Gaps: 1
US-09-782-953-3 (1-198) x US-09-644-868-7429 (1-1615)
Qy 1 MetGluGluValAspLeuGlnAspLeuProSerAlaThrIleAlaCysHisLeuAspPro 20
|||
Db 154 ATGAGGAGGTGGACCTGACGACCTGCCAGGACCTGCCAGGCCACCATCGCTGTCCCTGGACCCG 213
|||
Qy 21 ArgValPheValAspGlyLeuCysArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40
|||
Db 214 CCGCGTGTTCGTGGAGCGGCTGTGCGGGGCAAAATTTGAGTCCCTCTTTAGACGTATGAC 273
|||
Qy 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgVal-ArgIleAsnPheSerAs 60
|||
Db 274 AAGGACATCACCTTTTCAGTATTTTAAGAGCTTCAAAACGAGTCTCGAATAAATCTCAGCA 333
|||
```

QY 60 nProleuser-Ala1aAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLys 80
 DB 334 CCCCTTCTCCCGCAGAGTCCGAGCTCCAGCTGCTAAGACTGAGTTCTGGGAAAG 393
 QY 80 lUmetLysLeuTyPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProPro 100
 DB 394 AATGAAATTATATTTTCTTCAGACCTTACACATAGGAAGCTCACACTGGCTCCGCCAA 453
 QY 100 snProAspLysGlnPheLeuIleSerProProAlaSerProProValGlyTrpLysGln 120
 DB 454 ATCCAGACAGCAGATTCTGATCTCCCTCCCTCCCTCCGACAGTGGATGGAAACAG 513
 QY 120 alGluAspAlaThrProValIleAsnTyPheLeuLeuTyPheAlaIleSerLysLeuGly 140
 DB 514 TGAAGATGGACCCCGAGTAACTATGATCTTATATGTCATCTCCAACTCCAACTGGGCG 573
 QY 140 roGlyGluLysTyPheGluLeuHisAlaAlaThrAspProThrProSerValValHisIsv 160
 DB 574 CAGGGAAAAGTATGATTCACGCGCAGCTGACACACTCCAGCTGGTGTCCATG 633
 QY 160 alCyGluSerAspGlnGluAsnGluGluGluGluGluMetGluArgMetLysArgp 180
 DB 634 TATGTAGAGTATCAAGAG--AAGGAGAGAGAGAGAAATGGAAATGAGAGAGAC 690
 QY 180 roLysProLysIleIleGlnThrArgArgProGluTyPheProIleHisLeuSer 198
 DB 691 CTAAAGCCAAAATTTATCCAGACCGAGAGCGGAGTACAGCGCATCCACTCAGC 746

RESULT 9 US-09-644-869-8165

; Sequence 8165, Application US/09644869
 ; GENERAL INFORMATION:
 ; APPLICANT: Shyjan, Andrew W.
 ; APPLICANT: McCarthy, Sean A.
 ; APPLICANT: Holtzman, Douglas A.
 ; APPLICANT: Monahan, John
 ; APPLICANT: Richardson, Jennifer
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
 ; TITLE OF INVENTION: THEREFOR
 ; FILE REFERENCE: 1600.1182-001
 ; CURRENT APPLICATION NUMBER: US/09/644,869
 ; CURRENT FILING DATE: 2000-08-28
 ; PRIOR APPLICATION NUMBER: 60/151,062
 ; PRIOR FILING DATE: 1999-08-27
 ; NUMBER OF SEQ ID NOS: 9708
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 8165
 ; LENGTH: 1615
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(1615)
 ; OTHER INFORMATION: n = A,T,C or G
 ; US-09-644-869-8165

Alignment Scores:

Pred. No.: 3,04e-86 Length: 1615
 Score: 953.50 Matches: 188
 Percent Similarity: 95.50% Conservative: 3
 Best Local Similarity: 94.00% Mismatches: 6
 Query Match: 90.98% Indels: 3
 DB: 25 Gaps: 1

US-09-782-953-3 (1-198) x US-09-644-869-8165 (1-1615)

QY 1 MetGluGluValAspLeuGlnAspLeuProSerAlaThrIleAlaCyHisLeuAspPro 20
 DB 154 ATGAGAGAGGTGAGCTGACGAGCTGCGCCAGCGCACCATCGCTGTCNCCTGAGACCG 213
 QY 21 ArgValPheValAspGlyLeuCyHisArgAlaLysPheGluSerLeuPheArgThrTyAsp 40
 |||||

DB 214 CGCTGTTCGTGAGACGGCTGTGCCGGCCAAATTTGAGTCCCTTTANAGCATGATAC 273
 QY 41 LysAspThrThrPheGlnTyPheLysSerPheLysValArgVal-ArgIleAsnPheSerAs 60
 DB 274 AAGACATCATCCTTTTCAGTNTTTTAAAGCTTCAAAACAGTCTCGMAATTAATTCAGCAA 333
 QY 60 nProleuser-Ala1aAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLys 80
 DB 334 CCCCTTCTCCCGCAGACATGCGAGCTCCAGCTGCAATAACTGAGTTCTGGGAAAG 393
 QY 80 lUmetLysLeuTyPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProPro 100
 DB 394 AATGAAATTATATTTTCTTCAGACCTTACACATAGGAAGCTCACACTGGCTCCGCCAA 453
 QY 100 snProAspLysGlnPheLeuIleSerProProAlaSerProProValGlyTrpLysGln 120
 DB 454 ATCCAGACAGCAGTTTGTATCTCCCTCCCTCCCTCCGACAGTGGATGGAAACAG 513
 QY 120 alGluAspAlaThrProValIleAsnTyPheLeuLeuTyPheAlaIleSerLysLeuGly 140
 DB 514 TGAAGATGGACCCCGAGTAACTATGATCTTATATGCTCATCTCCAACTGGGCG 573
 QY 140 roGlyGluLysTyPheGluLeuHisAlaAlaThrAspProThrProSerValValHisIsv 160
 DB 574 CAGGGAAAAGTATGAAATTCACGCGCAGCTGACACACTCCAGCTGGTGTCCATG 633
 QY 160 alCyGluSerAspGlnGluAsnGluGluGluGluGluMetGluArgMetLysArgp 180
 DB 634 TATGTAGAGTATCAAGAG--AAGGAGAGAGAGAGAAATGGAAATGAGAGAGAC 690
 QY 180 roLysProLysIleIleGlnThrArgArgProGluTyPheProIleHisLeuSer 198
 DB 691 CTAAAGCCAAAATTTATCCAGACCGAGAGCGGAGTACAGCGCATCCACTCAGC 746

RESULT 10

US-09-644-871-7247
 ; Sequence 7247, Application US/09644871
 ; GENERAL INFORMATION:
 ; APPLICANT: Holtzman, Douglas A.
 ; APPLICANT: Pan, Yang
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
 ; TITLE OF INVENTION: THEREFOR
 ; FILE REFERENCE: 1600.1167-001
 ; CURRENT APPLICATION NUMBER: US/09/644,871
 ; CURRENT FILING DATE: 2000-08-28
 ; PRIOR APPLICATION NUMBER: 60/151,059
 ; PRIOR FILING DATE: 1999-08-27
 ; NUMBER OF SEQ ID NOS: 9739
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 7247
 ; LENGTH: 1615
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(1615)
 ; OTHER INFORMATION: n = A,T,C or G
 ; US-09-644-871-7247

Alignment Scores:

Pred. No.: 3,04e-86 Length: 1615
 Score: 953.50 Matches: 188
 Percent Similarity: 95.50% Conservative: 3
 Best Local Similarity: 94.00% Mismatches: 6
 Query Match: 90.98% Indels: 3
 DB: 25 Gaps: 1

US-09-782-953-3 (1-198) x US-09-644-871-7247 (1-1615)

QY 1 MetGluGluValAspLeuGlnAspLeuProSerAlaThrIleAlaCyHisLeuAspPro 20
 DB 154 ATGAGAGAGGTGAGCTGACGAGCTGCGCCAGCGCACCATCGCTGTCNCCTGAGACCG 213
 |||||

```
QY 21 ArgValPheValAspGlyLeuCysArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40
|||
DB 214 CGCGTGTTCGTGAGCGCCCTGTCGGGGCCAAATTTGAGTCCCTCTTTANGACGTATGAC 273
|||
QY 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgVal-ArgIleAsnPheSerAs 60
|||
DB 274 AAGGACATCACCTTTGAGTNTTTAAGAGCTTCAACCGAGCTTCGAATAAATCTTCAGCAA 333
|||
QY 60 nProLeuSer-AlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysG 80
|||
DB 334 CCCCTTCTCCCGCAGCAGATGCCAGGCTCCAGCTGCATAAGACTGAGTTTCTGGGAAAGG 393
|||
QY 80 LuMetLysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProProA 100
|||
DB 394 AATGAGATTATATTTGCTCAGACCTTACACATAGGAAGCTCACACCTGGCTCCGCCAA 453
|||
QY 100 enProAspLysGlnPheLeuLysSerProProAlaSerProProValGlyTrpLysGlnV 120
|||
DB 454 ATCCAGACAAGCAGTTTCTGATCTCCCTCCCGCTCTCCCGCAGTGGGATGGAACAAG 513
|||
QY 120 alGluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyP 140
|||
DB 514 TGGAGATGCCACCCAGTATGATCTCTTATATGCACTCCCAAGCTGGGCTCCATG 573
|||
QY 140 roGlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisV 160
|||
DB 574 CAGGGGAAAAGTATGAATTCACGACGAGCTGACACCACTCCCGCAGTGGTGTCCATG 633
|||
QY 160 alCysGluSerAspGlnGluAsnGluGluGluGluMetGluArgMetLysArgP 180
|||
DB 634 TATGTGAGTGTATCAAGAG---AAGGAGGAAGAAGAGGAATGGAAGAATGAGGAGAC 690
|||
QY 180 roLysProLysIleIleGlnThrArgArgProGluTyrThrProIleHisLeuSer 198
|||
DB 691 CTAAGCCAAAATTTATCCAGACAGGAGCCGCGAGTACACGCCGATCCACCTCAGC 746
|||
```

RESULT 11

```
US-09-652-123-7916
; Sequence 7916, Application US/09652123
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.1186-001
; CURRENT APPLICATION NUMBER: US/09/652.123
; CURRENT FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/151,135
; PRIOR FILING DATE: 1999-08-30
; NUMBER OF SEQ ID NOS: 9796
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7916
; LENGTH: 1615
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(1615)
; OTHER INFORMATION: n = A,T,C or G
US-09-652-123-7916
```

```
Alignment Scores:
Pred. No.: 3,04e-86 Length: 1615
Score: 953.50 Matches: 188
Percent Similarity: 95.50% Conservative: 3
Best Local Similarity: 94.00% Mismatches: 6
Query Match: 90.98% Indels: 3
DB: 25 Gaps: 1
```

US-09-782-953-3 (1-198) x US-09-652-123-7916 (1-1615)

```
QY 1 MetGluGluValAspLeuGlnAspLeuProSerAlaThrIleAlaCysHisLeuAspPro 20
|||
DB 154 ATGGAGGAGTGGACCTTCGAGGACCTGCGCCAGCCACCATCGCTGTCCCTGACCCG 213
|||
```

```
QY 21 ArgValPheValAspGlyLeuCysArgAlaLysPheGluSerLeuPheArgThrTyrAsp 40
|||
DB 214 CGCGTGTTCGTGAGCGCCCTGTCGGGGCCAAATTTGAGTCCCTCTTTANGACGTATGAC 273
|||
QY 41 LysAspThrThrPheGlnTyrPheLysSerPheLysArgVal-ArgIleAsnPheSerAs 60
|||
DB 274 AAGGACATCACCTTTGAGTNTTTAAGAGCTTCAACCGAGCTTCGAATAAATCTTCAGCAA 333
|||
QY 60 nProLeuSer-AlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysG 80
|||
DB 334 CCCCTTCTCCCGCAGCAGATGCCAGGCTCCAGCTGCATAAGACTGAGTTTCTGGGAAAGG 393
|||
QY 80 LuMetLysLeuTyrPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProProA 100
|||
DB 394 AATGAGATTATATTTGCTCAGACCTTACACATAGGAAGCTCACACCTGGCTCCGCCAA 453
|||
QY 100 enProAspLysGlnPheLeuLysSerProProAlaSerProProValGlyTrpLysGlnV 120
|||
DB 454 ATCCAGACAAGCAGTTTCTGATCTCCCTCCCGCTCTCCCGCAGTGGGATGGAACAAG 513
|||
QY 120 alGluAspAlaThrProValIleAsnTyrAspLeuLeuTyrAlaIleSerLysLeuGlyP 140
|||
DB 514 TGGAGATGCCACCCAGTATGATCTCTTATATGCACTCCCAAGCTGGGCTCCATG 573
|||
QY 140 roGlyGluLysTyrGluLeuHisAlaAlaThrAspProThrProSerValValHisV 160
|||
DB 574 CAGGGGAAAAGTATGAATTCACGACGAGCTGACACCACTCCCGCAGTGGTGTCCATG 633
|||
QY 160 alCysGluSerAspGlnGluAsnGluGluGluGluMetGluArgMetLysArgP 180
|||
DB 634 TATGTGAGTGTATCAAGAG---AAGGAGGAAGAAGAGGAATGGAAGAATGAGGAGAC 690
|||
QY 180 roLysProLysIleIleGlnThrArgArgProGluTyrThrProIleHisLeuSer 198
|||
DB 691 CTAAGCCAAAATTTATCCAGACAGGAGCCGCGAGTACACGCCGATCCACCTCAGC 746
|||
```

RESULT 12

```
US-09-652-915-9010
; Sequence 9010, Application US/09652915
; GENERAL INFORMATION:
; APPLICANT: Falb, Dean R.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.1175-001
; CURRENT APPLICATION NUMBER: US/09/652.915
; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 60/152,110
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 10311
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9010
; LENGTH: 1615
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(1615)
; OTHER INFORMATION: n = A,T,C or G
US-09-652-915-9010
```

```
Alignment Scores:
Pred. No.: 3,04e-86 Length: 1615
Score: 953.50 Matches: 188
Percent Similarity: 95.50% Conservative: 3
Best Local Similarity: 94.00% Mismatches: 6
Query Match: 90.98% Indels: 3
DB: 25 Gaps: 1
```

US-09-782-953-3 (1-198) x US-09-652-915-9010 (1-1615)

```
QY 1 MetGluGluValAspLeuGlnAspLeuProSerAlaThrIleAlaCysHisLeuAspPro 20
|||
```

```

Db 154 ATGAGAGAGGTGACCTCGACGAGACTCTCCAGCCGACATCGCTCTCNCCTGGAGCCG 213
Qy 21 ArgValPheValAspGlyLeuCyArgAlaLysPheGluSerLeuPheArgThrTyAsp 40
Db 214 CGCGTGTTCGTGGACGGCGCTGTGCGGGCCAAATTGAGTCCCTCTTANGAGGTATGAC 273
Qy 41 LysAspThrThrPheGlnTyPheLysSerPheLysArgVal-ArgIleAsnPheSerAsn 60
Db 274 AAGGACATCACTTCCTGAGTCTTTTAAAGCTTCAACGAGCTTCGAATAAATTCAAGCAA 333
Qy 60 nProLeuSer-ALAAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysG 80
Db 334 CCCCTTCTCCCGACGACGATGCGAGCTCCAGCTGCAATAAGACTGAGTTCTGGGAAAGG 393
Qy 80 MetLysLeuTyPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProProAla 100
Db 394 AATGAAAGTTATATTTTGTCTCAGACCTTACACATAGGAAGCTCACACCTGGCTCCGCCAA 453
Qy 100 nProAspLysGlnPheLeuIleSerProProAlaSerProProValGlyTyrPlyGlnTy 120
Db 454 ATCCAGCAAGCAGGTTCATCTCCCTCCCTCCCTCCGCGCAGTGGGATGGAACAAG 513
Qy 120 AlGluAspAlaThrProValIleAsnTyPheLeuLeuTyPheAlaIleSerLysLeuGlyP 140
Db 514 TGGAGATGCGACCCGACGATCAATACTATGATCTTTATATGCCATCTCCAAAGGGGCG 573
Qy 140 rGcLysLysTyPheGluLeuHisAlaAlaThrAspProThrProSerValValIleHisV 160
Db 574 CAGGGGAAAAAGTATGAAATGCGACGCGACTGACACACTCCGACGGTGGTGCATG 633
Qy 160 AlCyGluSerAspGlnGluAsnGluGluGluGluMetGluAsnGlyAsnGlyP 180
Db 634 TATGTAGAGTGATCAAGAG--AAGGAGGAAAGAGAAATGGAAMAATGAGGAGAGAC 690
Qy 180 rLysProLysIleIleGlnThrArgArgProGluTyThrProIleHisLeuSer 198
Db 691 CTAAAGCAAAATATATCCAGACGAGAGGCGGAGATACCGCGATCCACTCCAGC 746

```

RESULT 13

```

US-09-782-953-5
; Sequence 5, Application US/09782953
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674P21
; CURRENT APPLICATION NUMBER: US/09/782,953
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 5
; LENGTH: 597
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(594)
US-09-782-953-5

```

Alignment Scores:

```

Pred. No.: 1,99e-83 Length: 597
Score: 920.50 Matches: 178
Percent Similarity: 92.42% Conservative: 5
Best Local Similarity: 89.90% Mismatches: 12
Query Match: 87.83% Indels: 3
DB: 30 Gaps: 1

```

US-09-782-953-3 (1-198) x US-09-782-953-5 (1-597)

```

Qy 4 ValAspLeuGlnAspLeuPro-----SerIleThrIleAlaCyGHisLeuAspPro 20
Db 1 ATGATTTTATAGGACCTTTAGCTATACAAATTTAGCTCCGATTTGCTGTGGCAAAAGAT 60
Qy 21 ArgValPheValAspGlyLeuCyArgAlaLysPheGluSerLeuPheArgThrTyAsp 40
Db 61 GATGCTTCAGCGAAAGAGACCGGCGCAAAATTGAAATCCCTCTTCAGAAACATATGAC 120
Qy 41 LysAspThrThrPheGlnTyPheLysSerPheLysArgVal-ArgIleAsnPheSerAsn 60
Db 121 AAGGACACCACTTCACATATTTTAAAGCTTCAACGTCGCCGATTAACCTTCAGCAAC 180
Qy 61 ProLeuSerAlaAlaAspAlaArgLeuArgLeuHisLysThrGluPheLeuGlyLysG 80
Db 181 CCTTATCTGACGCGAGCGAGTCCAGGCTGGCTGCACAAAGCGAGTTCTGGGAAAGGAA 240
Qy 81 MetLysLeuTyPheAlaGlnThrLeuHisIleGlySerSerHisLeuAlaProProAsn 100
Db 241 ATGAAAGTTGATTTTGTCTCAGACCTTACACATAGGAAGTTACACACTGGCTCCGCCAAT 300
Qy 101 ProAspLysGlnPheLeuIleSerProProAlaSerProProValGlyTyrPlyGlnTy 120
Db 301 CCCGACAAACGTTCTCATCTCCCTCCGCTCTCTCTCCCTGCTGGGAAACAAGTA 360
Qy 121 GluAspAlaThrProValIleAsnTyPheLeuLeuTyPheAlaIleSerLysLeuGlyP 140
Db 361 GAAGATGCCACCCCGCATATAATTAGATCTTTATATGCCATCTCCAAAGGGGCGCA 420
Qy 141 GlyLysLysTyPheGluLeuHisAlaAlaThrAspProThrProSerValValIleHisV 160
Db 421 GGAGAGAAAGTATGAACTGCGATGCGACGACAGACCCCACTCCAGTGTGGTCCACGTG 480
Qy 161 CysGluSerAspGlnGluAsnGluGluGluGluMetGluAsnGlyAsnGlyP 180
Db 481 TGTAGATGACCAAGAGAAATGAGAGAGAAAGAGAAATGAGAGAAATGAGAGAGACCC 540
Qy 181 LysProLysIleIleGlnThrArgArgProGluTyThrProIleHisLeuSer 198
Db 541 AAGCCCAAAATCATCCAGACCGGAGACCGGAGATACCGAGATCCACTCCATTAGC 594

```

RESULT 14

```

US-09-782-953-11
; Sequence 11, Application US/09782953
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, R. SANDERS
; APPLICANT: ROTHERMEL, BEVERLY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO MUSCLE SELECTIVE
; FILE REFERENCE: UTSD:674P21
; CURRENT APPLICATION NUMBER: US/09/782,953
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/216,601
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 11
; LENGTH: 2331
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (144)..(734)
US-09-782-953-11

```

Alignment Scores:

```

Pred. No.: 8,16e-78 Length: 2331
Score: 872.50 Matches: 168
Percent Similarity: 93.09% Conservative: 7
Best Local Similarity: 89.36% Mismatches: 12
Query Match: 83.25% Indels: 1
DB: 30 Gaps: 1

```

US-09-782-953-3 (1-198) x US-09-782-953-11 (1-2331)

THIS PAGE BLANK (use 78)